A CRYPTOGRAPHIC SERIES

58

SOLVING CIPHER PROBLEMS

Cryptanalysis, Probabilities and Diagnostics

- ❑ Monoalphabetic Systems
- ❑ Polyalphabetic Systems
- ❑ Digraphic Systems
- ❑ Fractionating Systems
- ❑ Transpositions
- ❑ Probabilities and Statistics
- ❑ Solving Squares and Alphabets
- ❑ Diagnostics
- ❑ Problems to Solve
- ❑ Index

Historical Photographs

by
Frank W. Lewis

From Aegean Park Press

005.82 LEW
LEWIS, FRANK L.
SOLVING CIPHER PROBLEMS
OBC
003352200425

EM

s and Diagnostics

SOLVING CIPHER PROBLEMS

Cryptanalysis, Probabilities and Diagnostics

מכללת אורט בראודה
ספריה
קרית המכללה 78 כרמיאל 21982
מס' פרים 33572-000-001
מס' מיון 94 ס"ק
LEW

by

Frank W. Lewis

Copyright © 1992 AEGEAN PARK PRESS

ISBN: 0-89412-178-2 (soft cover)

AEGEAN PARK PRESS
P.O. Box 2837
Laguna Hills, California 92654
(714) 586-8811

Manufactured in the United States of America

FOREWORD

Because this is an appropriate spot in any book to make a few acknowledging bows to the various people or organizations who have contributed to the hoped-for success in the venture, I will take the opportunity to combine such "thank you" with an "object lesson."

I often bemoan the fact that most people could take on a lot more challenges than they usually settle for, in such areas as tackling an unknown problem, or learning bridge, or starting to learn Japanese, or daring to pick up the London *Times* crossword for a look-see. Any readers who feel a blushful tinge of self-recognition at my statement may be heartened by the fact that I am singling out my wife, Sylvia, as the first recipient of my gratitude for help in the present writing.

When we were snugly ensconced in our Maryland home, one might expect to find me dashing off to rehearse a chorus while she fulfilled a chore as president of the Greenbelt Women's Club. Or Sylvia would be having her picture taken with the children, as Mrs. COOP of the year while I (purely coincidentally) fulfilled my role as one of the directors at a board meeting of the same institution. But you get the idea. We just did different things together.

For example, my lunchtime-with-cronies was spent in scrambling through five or six London *Times* or *Telegraph* puzzle clip-outs, with occasional interspersing of drafts of my *The Nation* puzzles for proofing by the sandwich-munchers. But although Sylvia may have been pleased to observe the byline, as far as I know she never tried to solve my puzzles — or any other crossword — with any serious enthusiasm. And in keeping with security regulations, her knowledge of what I did at work was exceedingly skimpy, and her interest or understanding of such things as "ciphers" or "cryptograms" was non-existent.

Suddenly: retirement on the tiny, albeit Edenish island. Who to turn to for a guinea pig who could make sure that my next *The Nation* puzzle had no egregious errors? No logical candidate appeared among the local population or the relatively small number of expatriate friends who had sought refuge here. (Most of the latter were successful businessmen and wives, or people with practical or financial interests and pursuits — with the charming local golf course a number one interest, of course.)

When I joined the ACA and *The Cryptogram* began arriving every other month, it was probably my 13-year-old daughter, Wendy, who immediately showed aptitude and interest in solving the not-so-simple problems therein — including the foreign language ones which Wendy particularly liked.

To shorten the length of my story, by a combination of cajolery and desperate pleas for help, Sylvia was drafted willy-nilly into the puzzle-proving business, and since Wendy soon progressed to higher educational pursuits in England, Sylvia became more and more involved in my cryptanalyst hobbies, and eventually started supplying *The Cryptogram* with original problems-to-solve to a greater extent than I did.

While I enjoy the "better" crossword challenges we receive via bundles of Penny Press and Dell deliveries, it is Sylvia who orders the Logic Only numbers. What time she can spare from reading "better things," she spends on all sorts of crosswords and other puzzles, and she is getting darned good at it.

So now every single one of my weekly puzzles gets a thorough scrubbing by the lady of the establishment. And when the early "Mainly Crossword" volume of my Problem Solving was in production, and when this present book of Cryptanalytic Problem Solving was in draft form, Sylvia struggled through solution after solution of the many examples and can-you-solve-this items.

Thank you, Sylvia.

I should not forget such people as Paul Derthick, a former work-mate and author of the *Headline Puzzles* I use herein. My son, Gene Lewis, switched from mathematics to music at Oxford, but can still acknowledgeably keep a critical eye peeled for faulty thinking in my references to *Probability and Statistics*. My daughter Elaine, who acts as a stateside representative, will be happy to see that she doesn't have to go through the headaches of printing, storage and distribution of this present book, as she has with the more modest *Part One of Problem Solving*, and the reprint books of my puzzles which have periodically appeared.

The American Cryptogram Association several years ago gave a whole-hearted blessing to my using some of the problems which appeared in their magazine, as well as my adoption of much of the nomenclature. I express many thanks and trust that this book may inspire some of its members to continue the onward and upward procession.

Other people, such a Walter Penney, were kind enough to read the manuscript, and offer cogent suggestions and criticisms.

Naturally, if Wayne Barker, the guiding force behind Aegean Park Press, had not given his complete support to this effort, there would be no book. The countless phone calls from California to the Caribbean will somehow prove justified, one hopes, and the difficulties of long-distance publishing soon forgotten.

And special thanks to my *The Nation* fans, and the many friends we have made within the ACA organization. For when I receive an unusually encouraging letter from a puzzle solver (with quite a few such being relayed by either *The Nation* or Elaine, when fans order reprints) or an appreciative comment from one of the ACA KREWE who heard me babbling on as a convention banquet speaker, I realize that such gestures mean more to me than any monetary gain or other reward I might get from successful publication.

February 1992

—Frank W. Lewis

TABLE OF CONTENTS

I	Introduction	1
II	Elementary Cryptanalysis, General Remarks	5
III	Monoalphabetic Substitution (With Word Lengths Given)	11
IV	Monoalphabetic Substitution (Without Word Separators)	33
V	Polyalphabetic Systems	43
VI	Digraphic Systems	63
VII	Other Fractionating Systems	87
VIII	Other Ciphers	113
IX	Transposition	129
X	Probability and Statistics	149
XI	Solving Squares and Alphabets	163
XII	Diagnostics	175
XIII	Miscellaneous Problems	197
XIV	Problems to Solve	219
XV	Notes on Problems to Solve	227
	Index	247
	Bibliography	253

Chapter I

INTRODUCTION

SOME time ago, I prepared a lengthy manuscript on *Problem Solving* intending to cover a wide field, ranging from many types of crosswords and related puzzles to advanced cryptanalysis. It quickly became apparent that my reach was far too short to accomplish the projected stretch. For one thing, the area of corresponding interest between the many, many crossword solvers (including my modest contingent of *The Nation* fans)[1] and the relatively few cryptography buffs is small indeed.

Some amateur cryptanalytic enthusiasts, exemplified by those in the American Cryptogram Association[2] have specifically eschewed any reference to or inclusion of crossword puzzles in their publications. On the other hand, you seldom see reference to more than the simplest of cryptographic teasers in any American puzzle compendium. No Playfairs or other digraphic substitutions, despite the fact that such may be a critical ingredient in certain British crosswords of the more advanced type. In short, a crossword gobbler may well consider even moderately difficult cryptograms indigestible, and so-called cryppies may not like crossword puzzles in the least. What one group buys, the other has a tendency to ignore.

Another obstacle consists of the fact that there are at least three types of interest in cryptology which are rather disparate. The first is represented by the old-fashioned cryppie who loves to solve problems of medium difficulty in so-called hand systems: historic types such as Playfair and similar methods of communications used in World War I or the Civil War or earlier. Computers are largely eschewed, and pencil-pushing is at a premium.

Today a second type is making increasing appearance in the ranks of such organizations or interest groups as profess cryptanalysis for a hobby. This group consists of the computer buffs, who are not content to consider machine assistance as merely an honest method of disposing of the drudgery connected with writing things out, taking frequency counts, preparing charts, etc. Such experts derive their cryptographic satisfaction by creating powerful programs which will take a fairly difficult cryptogram and either (quickly and painlessly) serially follow the steps a pencil-and-paper solver might go through, or else figuratively beat it to death. Quite often, the interest in Playfair, or whatever, is minimal and perhaps the expert has really never tackled many cryptograms using the traditional approach. It may be merely a matter of "can I figure out a way of putting this cipher into my computer and having plain text spew out from the printer in record time?" Naturally, computer science has opened the way to many solutions which a hand-cryppie might think about if all the time in the world was available, and the rewards were sufficiently tantalizing. But such things as exhaustive trials are not all that much fun for the solver bent on outwitting the creator of a cryptogram.

A third type of professional interest comes from the growing body of communications experts who are concerned with the mechanics of high-speed keying

systems, transmission security, almost infinitely long streams of random key, and such.[3] Anything that happened in the field of cryptography last month may be exciting reading, but a tale of trench warfare concerning the reading of an ADFGVX message would attract little attention and possibly small comprehension.

There may be even a fourth body of readers which warrants consideration. Many books have been written on the general field of wartime "intelligence" — in particular that derived from the reading of enemy communications. Armchair warriors seem to derive pleasure from reading that the breaking of the German ENIGMA machine materially shortened the length of the conflict, even though the mysteries of the device itself might be considered completely arcane. Can one book on *Problem Solving* offer anything of interest to all these widely ranging tastes?

The answer seemed to be to publish a rather modest First Part of the *magnum opus*, to include British and American crosswords of many types (ranging from kid stuff to very challenging brainteasers) as well as chapters on other word games and non-crossword puzzles. The field of cryptography was tackled in a minor way, mainly because many publications now offer simple monoalphabetic substitutions to their readers. The intent was to test the waters, and possibly publish a much larger *Part II* adjunct which would range much further and more completely into the areas of more advanced cryptanalysis, and include chapters on Diagnosis, Elementary Probabilities and Statistics, and perhaps Problem Solving in General.

Reading between the lines, you can see the missionary aspect of my endeavors to entice the run-of-the-mill crossword solver into the more rarified atmosphere of both the harder type of crossword and the kind of cryptogram which is a cut above the very simple substitutions they find in popular publications. And, obviously, I sneak in a good word for the better crosswords whenever I have the pleasure of talking to or with my cryppie buddies. I leave it to the reader to decide whether or not my proposed Problem Solving Duo can be forced into adjacent space on the book shelf, and herewith offer the follow-up to *Part I* — which was labeled *Problem Solving, with Particular Reference to the Cryptic (or 'British') Crossword and Other 'American' Puzzles.*[4]

NOTES ON INTRODUCTION

(1) The author has been the perpetrator of the fairly difficult *The Nation* puzzles since 1947. They are similar to the British (or "cryptic") crosswords exemplified by the London *Times*. Although for many years this type of puzzle was a maverick over here, and considered "too hard" by most publications, it has now reached a stage of acceptance that has caused a plethora of Cryptics in many magazines — varying in quality from awkward to highly professional. And more and more "setters" are getting into the act.

(2) The American Cryptogram Association is one of the most active amateur solving groups, of sixty years standing. Their official publication, *The Cryptogram*, appears every other month, and contains five departments covering Aristocrats and Patristocrats (monoalphabetic substitutions), Cryptarithms (in which digits of mathematical problems are replaced by letters), Xenocrypts (problems of varying difficulty in many languages, including artificial ones such as Esperanto) and a Cipher Exchange containing more difficult systems, such as we will meet later on in this book. Most of the problems are labeled as to type, such as Playfair, but some are listed as "unknown" and a special section is included as an Analyst's Corner where the solver may or may not be given a hint as to the proper classification. There are also articles written by members on historical systems, and how-to, as well as book reviews of relevant publications, a Computer Column, and general organizational information. Queries concerning the organization and the *Cryptogram* may be sent to the ACA Treasurer, 18789 West Hickory St., Mundelein, IL, 60060.

(3) A good example of a technical publication which includes among its readers a large group of communication experts is *Cryptologia*, published at Rose-Hulman Institute of Technology, Terre Haute, IN 47803. Many of the articles have titles which point in the direction of Data Encryption Standards, mathematical and mechanical methods, ASCII encipherment, authentication and such. There are also a few articles concerning more traditional cryptanalytic matters, such as successes in WW-II, or the challenge of some cipher found in the British Museum, or some tidbit about the giants in my field, such as Friedman, or Yardley, or Tiltman, or Turing.

(4) For those who may wish to read *Part I*, covering many types of puzzles and simple cryptograms, they may get a postpaid copy by sending $8.95 to E. S. Lewis, 125 Quaker Lane, Rochester, MA., 02770. Reprint books of representative *The Nation* puzzles are also available from Ms. Lewis.

Chapter II

ELEMENTARY CRYPTANALYSIS
General Remarks

THERE is only one technical field I know of in which most how-to books available to the public are necessarily written by people, with certain exceptions, not professionally engaged in that science or technology. Surgeons can tell you how to cut, artists can tell you how to hold a brush, and writers can even tell you how to write. But if you see a book entitled *Secret Communications of the Brownshirts* or *How I Decoded Hitler's Invasion Plans*, you can be almost certain that the author, especially if American, did not push a pencil in some Allied Cryptographic Bureau during the war.

It is true that a good deal of information is obtainable about PURPLE, the high-level Japanese cipher machine that the justly famous W. F. Friedman[1] "broke," naturally not single-handedly. And there are numerous accounts of the role of cryptanalysis in some of the Pacific naval battles. The so-called best-kept secret of the British successes with the ENIGMA machine (the gadget that many branches of the German high command trusted to convey their most "Strengst Geheim" affairs, including plans and operations) is no longer a secret. But Mr. Friedman did not write any of the stories you will read about PURPLE. You can thank (or blame) the Pearl Harbor investigation for that revelation. Naturally, the "Winds" messages,[2] possible hints of impending hostilities — the whole story of "were we caught napping?" — was of great importance at the time. Because a good deal of the tale had leaked out, it would have been stupid to keep PURPLE completely under wraps.

With some exceptions, even the plethora of information poured out in recent years concerning the work of the British at Bletchley Park on the German ENIGMA and — horrors — even more sophisticated machines was more likely than not to stem from someone associated with the technical operation. Excellent writers though they might be, many tell-all types were in the intelligence game as distributors, or users, or go-betweens, but might have only a cursory knowledge of how the pencils were actually pushed.[3]

In any event, and perhaps stemming from the Yardley[4] flap a half-century ago, most stories of American successes with wartime German or Japanese communication systems are the work of people removed from the inner core of the business. And even more innocuous cryptanalysis how-tos are likely to bear the names of diligent but non-professional students of the field, who know the classic systems, and can make and break moderately difficult ciphers, but who never saw an enemy cablegram.

The above effusion may seem out of place, considering the fact that this is NOT an account of my years of down-to-the-mat struggle with us-against-them wartime adversaries in the field of communication intelligence. Beyond casual references to ENIGMA or PURPLE or other relatively familiar machines, there will be little mention of the more exotic problems in real-life cryptography. Rather, we will talk about "fun"

systems, sometimes of minor historical importance, that are well within the scope of the average solver, without extensive machine aid or wide experience. It would have been satisfying to include more specific references to certain methods of transposition used by German agents, or twists on Playfair used by perhaps lesser developed countries, or the ins and outs of tackling a difficult-but-solvable code, or an exciting spy cipher. In truth, the kinds of systems I write about are by now almost entirely old hat, and would never be used by modern cryptographers worthy of that appellation. I must defer to our current authorities, who seem horrified by the thought that the attache of some fourth-ranked country, who has stumbled upon the cryptographic principle of some wartime system, might learn that his step-child is a blabbermouth if he reads how easy it is to solve.

So much for the professionals and their wartime anecdotes, which may stray into the no-no territory. However, some areas of opportunity exist for experienced analysts, regardless of their professional or non-professional background, to inculcate some of their specialized knowledge into a growing group of enthusiastic (though usually non-professional) cipher nuts. Naturally, any writer must use good judgment in demarcating the line between straightforward discussion of non-secret, well-established techniques and more lurid exposé treatment of matters that might curl the hair of our more conservative security watchdogs. After all, they're paid for being careful. If in doubt, get out the old blue pencil. One of these areas covers scholarly research into the mathematics underlying many phases of cryptanalysis, as well as communications in general. Computer science is another area. I have pointed out that many solvers seldom tackle even the more simple problems by hand and really don't have that much love for ciphers — they spend all their time figuring out machine applications both in Cryptography (the creating) and Cryptanalysis (the breaking).

In the first of these areas, one of the truly professional analysts, Abraham Sinkov, has published a book on *Elementary Cryptanalysis—a Mathematical Approach*.[5] If the reader gets impatient of any low-level arithmetic I may trot out in subsequent discussion of distributions, multiplication of matrices, etc., he should remember that I belong to the apples-and-oranges category of non-mathematical statistician who may have a practical sense of probabilities but shies away from a Sigma sign or a complex equation as being a member of the viper family. Better you turn to Sinkov for the real "gen," as the British say, if you insist on a higher-level mathematics explanation.

It doesn't sound like much, but the enthusiasts in the ACA who have read so much about Friedman, who collect memorabilia relating to cipher machines of WW-II, and who may have libraries of specialized material that a professional would envy, welcome with open arms any veteran of the establishment who may have talked with Yardley, or has actually been in Hut 6, or picked a foreign cipher machine to pieces. Therefore, they have said they don't mind if I use their publication, the *Cryptogram*, as a vehicle for describing several not-very-high-level but-sometimes-challenging systems they include in their manual. Since the book they hand out with membership details the method of construction, but NOT any suggested method of solution for problems within their repertoire, my scribbling may even be of some use to them.

What kinds of systems are we going to talk about now? First of all, Simple Substitution. In *Part I*, I touched lightly on the subject, covering certain Monoalphabets, and the way in which the enciphering components are used. But the more orderly classification of the varying degrees of sophistication established by ACA deserves much more extensive treatment. Then there is a wide assortment of Polyalphabetic Systems, which may employ the setting of plain and cipher components at many different settings, reducing greatly the opportunity for simple statistical solution.

Digraphic and Trigraphic systems of several kinds will be covered, and some of the many Transposition methods explored. Also, some mention will be made of other methods of making one's communications secure, and cooked-up problems which approach real-life situations may be presented. Finally, there must be some treatment of Diagnostics — in other words, how do we pick up a problem? And it goes without saying that we must have an understanding of elementary statistics and probability, in which fields I may require the assistance of more mathematically oriented advisors. We might touch on the subject of problem solving in general, and make a few references to the decisions that people have to make in their everyday lives, if they play cards, word games, or try to figure out the chances of such and such happening. So on with the dance, if Holmes' "Dancing Men" can serve as a cue for "Simple Substitution."

NOTES ON CHAPTER II

(1) Possibly the first name in the pantheon of eminent American cryptologists is that of William F. Friedman. Since he was the man who as head of the Signal Intelligence Section (under the Chief Signal Officer) hired me in 1939, I wish I could have written the book which Ronald Clark turned out in 1977. Since he died in 1969, Mr. Friedman could have hardly recounted his fabulous stories to Mr. Clark, who is an excellent professional biographer but hardly within the Friedman circle.

These are the stories which my wife and I heard in the good old days while hanging on to a glass at one of the office parties at the Friedman house. Since the entire organization numbered about forty, including junior clerks, a few enlisted men, the typing pool and so on, one can sense that many of the get-togethers needed little more than a large living room where some of us would be sitting on the sofa, or more likely the floor, listening to outrageous tales of the boss's experiences at Col. Fabyan's Riverbank laboratories in 1915 or so. In any event, you should read *The Man Who Broke Purple* by Ronald Clark, and also turn to the chapter on "Two Americans" in the excellent *The Codebreakers* by David Kahn. Incidentally, Kahn is a splendid example of the type of cryptologist who was "not in the business" with the Government, but by dint of proper research and plenty of digging knows more about the field and what must go on behind closed doors than most of the insiders.

(2) The so-called Winds Messages were a provision made by the Japanese to notify recipients of impending hostilities by an open code consisting of innocuous weather bulletins such as "east wind rain" — which meant Japan-U.S. relations were in danger. I know little about the transmission of such messages before December 7.

(3) There are several good examples of the writers and historians who were close to but not part of the communications-intelligence production line. As early as 1976, William Stevenson (*The Man Called Intrepid*), wrote a good deal about Bletchley Park, where the British deciphered the German ENIGMA messages. Intrepid was one of the giants in the intelligence game, but not a leading cryptanalyst. Other names who were either part of the intelligence distribution system or journalists or military historians or scientists come to mind — too many to list. But one immediately thinks of Anthony Cave Brown (*Bodyguard of Lies*) and Ronald Lewin (*The American Magic*) *inter alia*. As a matter of fact, I find myself consulting Lewin's tabulations when I try to remember how many people I had, when I was Civilian in Charge of a large section at Arlington Hall during the war. He can write about my section, but in theory I am forbidden to do so.

An exception to the rule that "those who know don't write and those who write may not know" is a good book on ENIGMA by Gordon Welchman, entitled *The Hut Six Story*. He not only was on the scene but he also played a major role in the breaking and exploitation of the system that yielded so much vital information to the British. For some reason the authorities over there have been lenient during the last decade in their decision to release many of the highly exciting tales of heroic accomplishment in the war of the ether. They may have been annoyed with Welchman, but the source of their wrath was his roaming into areas where his expertise may not be as well established as his cryptographic skill; the second half of his book gets into "today" discussions where he might be accused of telling present-day authorities how to run many railroads. I am reminded that perhaps too often we talk about "the British solution of ENIGMA" without giving due credit to the brilliant Polish cryptanalysts who did much of the ground-breaking in the early thirties. Without the work of the Poles to build upon, it is conceivable that the Hitler-crushing victories of the Bletchley crew would never have materialized—or at least time would have been irrevocably lost. A more prolific writer, Sir Harry Hinsley, is another exception to the but-you-weren't-there complaint, for he worked at Bletchley, and is the main author of the official *British Intelligence in the Second World War*. Would that the United States had an equivalent official story written by an on-the-scene reporter.

The leniency of the British may have stiffened under the aegis of Margaret Thatcher, particularly since the battle almost-royal between the government there and Peter Wright, the author of *Spycatcher*. A former MI wheel, the writer spilled a lot of embarrassing beans about the inner workings of the intelligence service. I usually argue that we should be a little more informative about the past, but my sympathies in this case extend a bit towards Maggie. Things are talked about which, while perhaps juicy tidbits for the reading public, could jeopardize present operations at worst, and chill relations between allies at best. Telling how best to bug your friendly neighbor's living room is not conducive to future backyard barbecues. But naming names and citing dates in the very recent past is quite different from explaining how the enemy's codes were broken in 1943. Lewin lists the date, but naturally doesn't know exactly how it was done. My argument is that after fifty years, we ought to be able to write about events that could not possibly jeopardize present operations.

(4) Herbert O. Yardley is the other name in Kahn's "Two Americans" chapter and usually gets twin billing with William Friedman. I met Yardley only after he was more or less over the hill, after his unsuccessful attempts to get back into the business with our government. But in his youth, and his prime, he was an awesome figure, deserving of the respect given him by all ranks of cryptanalysts. As a comparatively youthful and lowly clerk in the State Department code section, he shook the trees by casually announcing that he could read our official communications very readily (no pun intended) if he were to tackle them analytically.

Later on, after learning a thing or two from the French cryptanalysts during WW-I, he set up shop with a handful of people as a more or less clandestine operation in a New York City brown house, being funded by military and State Department sources. He had a field day with the most secret communications of many foreign offices, and the juiciest of his output found its way into a folder placed on the desk of the Secretary of State each morning. But in 1929, when Henry L. Stimson took over as Secretary, all hell broke loose on the first morning that the unexpected folder was opened. "Gentlemen do not read other people's mail!" — an apocryphal announcement, but you get the general idea.

Accordingly, the rug was pulled out from under the nimble feet of H. O. Y. With no continuing government slot to fit into, Yardley turned to writing, and churned out a series for the *Saturday Evening Post*. Eventually, in 1931, the fan was hit by one of the most explosive books ever published: *The American Black Chamber*. In it, Yardley recounted his highly successful attack on the communications of many countries. Of particular import was the story of his "reading" of the instructions given by the Japanese government to their representatives at the International Naval Conference of 1923, in which they were arguing for parity in the strength of battleships with England and the United States. France was also involved, but willing to settle for a slightly lower ratio. Briefly, the message coming through was "argue till the last minute for 5-5-5, but realistically we know you will have to settle for 5-5-3." And that is how the so-called 5-5-3 Treaty came into being — but you can imagine the kind of poker game that went on at the conference when we knew exactly what cards the Japanese players were holding. Naturally the book was an immediate best seller in Japan, and a source of much concern to the United States. The Japanese were properly outraged, and I along with many others feel that the seeds for anti-Americanism were sown that were harvested at Pearl Harbor.

Our government managed to "stop the presses" but total recall of this explosive book was impossible. A law was passed by Congress which still exists as a barrier to my ever divulging any information about or contained in "signal intelligence" — no time limit, no exceptions. For many years *The American Black Chamber* was a banned book. Naturally a few copies were hoarded by fortunate cryptanalysts, but for almost fifty years you could not purchase the book. I find the notations near the front of my copy of a Ballantine Books Edition very interesting — Copyright 1931, C. 1981 by Herbert O. Yardley. Most such notices have many dates, with only a few years between various printings.

(5) — Sinkov is one of the triumvirate hired by Friedman in the early thirties as the nucleus around which the Army's Signal Intelligence Section was built. Along with Dr. Solomon Kullback and Frank Rowlett, he continued a brilliant career until retirement a couple of decades ago. All three served as officers, getting their Colonel's eagles at exactly the same time. Sinkov was a senior member of the Australian and later Indian headquarters of our overseas signal intelligence effort, while the other two stayed home in command roles. Since I served directly under Kullback, and continued to be part of his team even while nominally working with Operations when he was head of R & D, I was naturally closer to him. But I worked enough with the other pair to entitle me to attest to the fact that this brilliant trio would stand up well in comparison with any selection of professionals in government or business.

Chapter III

MONOALPHABETIC SUBSTITUTION
(With Word Length Given)

THE simplest type of so-called Cryptogram one sees in newspapers, puzzle magazines and the ACA bimonthly is Monoalphabetic encipherment. This means that each distinct plain value is replaced by a unique cipher value. Thus every T in a plain message may have a Q in the corresponding position of the cipher, every H going to Y, and every E going to T, perhaps. Under certain conditions the plain and the cipher letters may be the same, A going to A.

There are several accepted ways in which this transformation may be performed according to rule. In most cases the encipherer is aided by reference to two sets of "alphabets" — the (usually) 26 letters being placed in normal order, or random order, or some prearranged order on each "strip" or row of consecutive cells on cross-section paper, or whatever the physical set-up dictates.

To "encipher" the message, one notes the first "plain" letter and finds it on the upper sequence, then replaces it with the "cipher" letter in the cell immediately beneath that spot on the lower sequence. One then takes the second plain letter, finds the proper "substitution", and so on through the entire message.

Enciphering clerks are quite likely to realize that time is saved by going through message and putting down a Q under every plain T (if that is the proper substitution), a Y under every H, and so on, just as a deciphering clerk may quickly go through all the Q's of the cipher message and replace them with plain T. But this is merely a matter of mechanics.

In the simple ciphers we are first investigating, the word length may be preserved. Thus, THE will be replaced by a 3-letter cipher equivalent, such as QYT — and it is of course an obvious weakness of such encipherment that every THE will result in the same QYT, giving the "guesser" quite a head-start on solution.

I will mention the several ways the plain and cipher alphabets can be prepared by the users according to prearranged plan, but we will leave the important discussion of how it affects solution, and alphabet reconstruction in general, to a separate section further on in this story.

For the time being, note these possibilities. One can employ NORMAL alphabets for both plain and cipher, lining up the A-Z sequences so that A is over D, B is over E, and so on. This results in what is termed a CAESAR slide — since that gentleman was supposed to have enciphered his communications by adding three to the value of the plain letters (in normal order, considering A to be 1, B to be 2, etc.) Thus HIC would come out as KLF if he had used our modern arrangement of 26 letters, which I'm sure he didn't.

This is rather too insecure, isn't it? No matter how the alphabets were "slid," NOT would yield a cipher with the first two letters of a three-letter word as a consecutive pair: FG-, or NO-, or whatever, with the third letter also arithmetically related.

Would it help if we put a NORMAL alphabet on top, and a REVERSE alphabet (running from Z back to A) on the bottom? This would give us plain A going to cipher Z, plain B going to Y, etc. Not much of an obstacle, and even a relatively inexperienced solver should spot such usage, if a good guess as to plain-text were forthcoming.

So let's try putting a NORMAL alphabet on top, at the same time constructing a cipher alphabet in some random or even prearranged way, and pairing the two so that A always came out D, perhaps, but B would be X, C would be G, and so on. That takes care of the constant mathematical relationship of NOT mentioned above, but still has a weakness which will be obvious when we discuss alphabet reconstruction later on.

We could also use a MIXED (in some prearranged way) alphabet on top and the same alphabet (at a different slide) on the bottom for Plain and Cipher pairs. Or, in an effort to make things a trifle tougher for the solver, we could use a MIXED alphabet for the plain and a DIFFERENT MIXED alphabet for the cipher.

These various set-ups are associated with numbers by the members of ACA, and for convenience I will borrow their terminology. Key 1 (abbreviated to K1 when associated with monoalphabets) designates the type in which the Plain is MIXED, and the Cipher NORMAL. (How the alphabet is MIXED is irrelevant at the moment.)

K2 implies that the Plain alphabet is NORMAL while the cipher strip is a MIXED alphabet. K3 involves a MIXED alphabet being slid against the same MIXED alphabet, which creates certain phenomena to be discussed later when we talk about solution. K4 uses two different MIXED alphabets.

In dealing with any of these problems there may or may not be present a "hint" (or crib) as to some word or phrase which appears in the puzzle. Often this "tip" will be slightly disguised in the form of a CAESAR slide, described above as a NORMAL alphabet being slid against a NORMAL sequence.

This has the effect of substituting each plain letter for an element removed on the normal sequence by the same amount. The "tip" can be recovered very easily by writing down the enciphered version given, and then extending each letter on a normal sequence:

Tip as it appears	**BKN**
Extended by NORMAL	**CLO**
	DMP
	ENQ
	FOR
	GPS
	etc.

A glance will reveal the FOR appearance as a "crib" to try. (Very rarely, in short words particularly, there may be two good-looking possibilities which have to be tested.) Computer buffs will of course have a convenient program for CAESAR

extension, so all one has to do is enter the cipher version, and scan the 25 rows of extension. Less privileged cryppies will still need pencil and cross-section paper. Now let's examine a cryptogram based on one of the simpler key types mentioned, K1.

Monoalphabetic Substitutions Using Normal Cipher and Mixed Plain Components (K-1)

```
*KNYNXN'A KILYMLKNO REKJIH LA PNYNYNA, NOHUJBTU MNMNJ NYQ
 NPNMN SLPRI (AJBIMR JS *XNYLON URXK) UNCR PRMJXR LXKJIHNYH.
```

The asterisk in the above indicates that the word is a proper noun. This example, since it is of the variety which ACA dubs "Aristocrat," very kindly maintains punctuation as well as word length. In practice, one could probably jump into the solution by finding out what happens when you call a likely three-letter grouping THE, or AND or FOR — or a two-letter word IN or TO (if the letters occur often enough) or OF (if the second letter is low frequency.)

Had we not been told, we might first ask "is this a simple substitution with original word divisions, or could it be some other cipher?" Casual inspection says that it probably is a mono, since some of the letters are very frequent, and there appears to be digraph cohesion and repetition.

Frequency by itself wouldn't be 100% conclusive, since transposition would have the same frequency "look" (though matching that of plain text) while repetitions of digraphs and longer would be considerably fewer.

Sometimes one can jump feet-first into the "probable-word" method of solution, and guess from some word pattern, or a suggestion about the content of the message, or the frequent appearance of a trigraph.

At other times, one has to more carefully analyze frequencies, taking counts and noting specific behavior of certain letters, etc.

In this case, the "title" of the cryptogram is "Agricultural Economy" and the asterisks should alert us to the fact that some country name may be present. Since the first word is "proper" and ends with an apostrophe plus "A", we can deduce that the A represents plain S, and the message might start with something like CANADA'S or PANAMA'S or any proper name that has the repeated pattern of N which is so obvious. In the two names suggested, we could fill in all plain A and N occurrences, and the plain S, so let us do so.

But one quickly notes that the fourth and fifth word would end with S, and the rhythm of the sentence goes "Somebody's something something IS .....S" — and the word after IS has a pattern that CANADA or PANAMA should make obvious.

```
*KNYNXN'A KILYMLKNO REKJIH LA PNYNYNA, NOHUJBTU MNMNJ NYQ
 CANADA'S C IN  CA         IS  ANANAS  A        A A  AN
 PANAMA'S P IN  PA         IS  ANANAS
```

```
NPNMN SLPRI (AJBIMR JS *XNYLON URXK) UNCR PRMJXR LXKJIHNYH.
A A A  I     S           AN  A   DC   A       D  IDC   AN
                             MP           M   MP
```

Admittedly the proper name could be something less likely. (Just for the fun of it, I tested the pattern on a computer program which throws out all pattern words, and out of over 30 occurrences of ABCBDB the "proper" names besides CANADA and PANAMA were BIKINI, BIMINI, HAVANA, MALAGA, OAXACA, RIMINI, SAHARA and TARAWA.

If you insist, you can try them all, but looking just at the Canada/Panama choice you will quickly spot the fact that Panama is the winner. Do you like a four-letter word ending in DC, or would you rather have MP?

And it takes little imagination to put in IMPORTANT for the last word.

```
*KNYNXN'A KILYMLKNO REKJIH LA PNYNYNA, NOHUJBTU MNMNJ NYQ
 PANAMA'S PRIN IPA    PORT IS  ANANAS  A T O     A AO AN

NPNMN SLPRI (AJBIMR JS *XNYLON URXK) UNCR PRMJXR LXKJIHNYH.
A A A  I  R  SO R   O   MAN  A   MP   A      OM  IMPORTANT
```

I am sure you would quickly fill in the entire message, on the basis of "principal export is bananas":

```
*KNYNXN'A KILYMLKNO REKJIH LA PNYNYNA, NOHUJBTU MNMNJ NYQ
 PANAMA'S PRINCIPAL EXPORT IS BANANAS  ALTHOUGH CACAO AND

NPNMN SLPRI (AJBIMR JS *XNYLON URXK) UNCR PRMJXR LXKJIHNYH.
ABACA FIBER (SOURCE OF  MANILA HEMP) HAVE BECOME IMPORTANT
```

It is common practice to have some rule of composition for an alphabet strip or sequence, rather than just scribble down a jumble of letters. For one thing (to make our story more exciting) it may involve a secret agent or some such cloak-and-dagger operation, wherein communicants must be able to remember little details such as keywords, in order not to get caught with a bag full of reference books and other paraphernalia.

If you read my earlier book you might remember that I showed a Playfair square associated with the nastier English puzzles. It was based on my name, so let's try ABRAHAM LINCOLN for a better known starting point. If we write down the letters as they occur, omitting any repeats, we get ABRHMLINCO, correct? Now add the remaining letters of the alphabet in order to get:

```
ABRHMLINCODEFGJKPQSTUVWXYZ
```

In the case at hand we are going to use the alphabet as a horizontal sequence rather than in a square. There are many more complex ways of generating a sequence in some prearranged way, and these may be discussed when we take up the matter of square-and-alphabet construction more specifically.

Let's look at the Plain Alphabet we get, when we put a normal sequence underneath the values obtained:

```
S U V   X     T R O P I C A L B   E F G H     M N
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
```

Obviously the keyword is TROPICAL, and the other values which do not appear could naturally be filled in at the proper spots.

Monoalphabetic Substitutions Using Normal Plain and Mixed Cipher Components (K-2)

The first K-2 example we will use goes:

```
AM QWL W QALF DWG AGZFFZ QNH HGYF LWAZ MNWM AG MNF FGZ
MNF BHPF THO MWEF AL FJOWB MH MNF BHPF THO DWEF.
```

It also has a Caesar tip of DPTZK. (Don't read any further without grabbing a pencil and extending the normal sequence down the letters of the tip, to find your "hint". Geniuses — or is it genii — will of course spot the correct word without recourse to pencil and paper, though this displacement is so close that the discovery may be considered trivial.)

Since there is only one 5-letter word in the whole message, FJOWB must be the location of the tip, and alphabets can be started — Normal on top, in keeping with the K2 rules.

```
Plain:   ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher:  W   F       B    J   O
```

Obviously F is in the keyword (probably begins it.) and the rest of the sequence runs from Plain K or L "around the corner" to C or D.

Jumping into the wild blue, one can make a stab at many words here. No wonder this was the first mono on the page labeled ARISTOCRATS, by which name the ACA recognizes monoalphabetic substitutions which are divided into word length. (Monos in which no word separation is given, with the solver being handed only a string of 5-letter cipher groups, are termed PATRISTOCRATS.)

Obviously the more flattering name implies a much more difficult lineage, and this type will be discussed further along. The five-letter groupings are for convenience in bookkeeping here (and a standard procedure in radio transmission) without reference to word endings and beginnings.

The 3-letter word ending in cipher O is undoubtedly YOU, which leads to the assumption of MH for TO. This leads to THAT and THE, and WHO, and the problem is half over. However we should recover the keyword to wrap up the solution.

```
Plain:   ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher:  W   F  N   B  HIJKLMOPQ T
```

You will of course have added cipher letters such as K and L merely because they close the sequence gap.

It is a relatively simple chore to recover the text "It was a wise man indeed who once" etc. And the components appear as:

```
Plain:   ABCDEFGHIJKLMNOPQRSTUVWXYZ
Cipher:  WXYZF  NA EBDGHIJKLMOPQ T
```

The Keyword F—NA-E must have a vowel in the second or third position, and the only candidate left is U. This puts V cipher as Z plain.

We have not placed C and one of the R-S pair, but the only sensible answer seems to be FURNACE, and since we have read the message and there is no more information, we will leave it at that.

Monoalphabetic Substitutions Using the Same Mixed Sequence for Both Plain and Cipher (Juxtaposed at Some Interval) (K-3)

Let's tackle the thing together, and talk about techniques as we come to them.

There is no "tip" for this one, and it may be moderately difficult, judging from its position in the series going from easy to nasty. There may be a "title" at the head of the puzzle, such as "Big business" or "The good old days" which sometimes helps you make a stab at a probable word. In this case the title is "Sticklers" which doesn't suggest much at present, but may help later.

We can't get much help ahead of time from the fact that it is a K3 example, except for one thing. In a monoalphabetic substitution using the same sequence for both Plain and Cipher, it is patent that no plain value can be represented by that cipher value.

(E cannot be E while all or some of the other values do not carry over from Plain to Cipher — and if one is the same, all must be the same —which would result in a real booboo mark on the cipher clerk's record.)

Thus, in the cipher text below, IM cannot decipher to IS or IN or AM:

```
DRVMELWXVJIQY LDHHWQC, DRUQORMNRCI IM JRVXQRI CMWFQSC,
ZSQBDQRIWG SQBDXSQ QKIQRCXFQ NMSO, IJK EMYQSR JUXWXIXQC.
```

I'm sure all potential solvers have a reasonably good idea about the fact that some letters (E, T, for example) are very common, some (such as L or H) hit a middle frequency, and some are very rare (Q, X and J for example.)

There have been several versions of the "accepted" frequency orders, such as the ETAOINSHRDLU associated with Line-o-type slugs. (The operator might run his hand across the keys to fill in a slug which had a mistake, and would later be yanked, and the keys on the machine had been placed in this relative order.)

The order arrived at is strongly dependent on the type of text one chooses. Newspaper text might yield a sample a little bit different from that of literary text, and military telegrams (which probably eliminate a lot of "the's" and such) may be completely different. It also makes a difference whether you are counting 1,000 or 50,000 letters.

The order I'm using in my computer program, based on a large sample of text, is as follows:

E T A O N R I S H L D C U P F M W Y B G V K Q X J Z

The percentages may vary from sample to sample, but usually (for English text) it runs from around 13% for E to less than 1% for the last several letters.

Our problem message is long enough to perhaps make statistics applicable, so let's take a "frequency count". In the simplest of cases, a modest "check and tally" might be sufficient; write down the normal alphabet, and put a mark below the letter corresponding to the cipher letter as you take each in turn, for the fifth occurrence crossing the four vertical strokes with a diagonal line:

```
A B C D .. etc.
- - - -
  - - -
    - -
    - -
```

But if the going promises to be rough, it may be worth taking a more informative count, showing the "running digraph" frequency (within words) by writing the following cipher letter under the position of the letter being considered, rather than just a stroke.

Since the frequency of beginning and ending letters is rather specialized, a small gimmick may be used by the counter which will possibly be very helpful. If you figure out some way of telling yourself that "this is a beginning letter" or "this is a final" you can spot the plain Y or plain S, perhaps, or say "this is probably not an E, even though the cipher has a high frequency."

One way would be to draw a circle around the second letter of any beginning digraph — circle the R under the D when you enter DR, but don't circle the V under R. And since there is no "following letter" for the last letter of a group, just put a vertical line ABOVE the cipher letter being counted when it happens to be in the last position:

```
    =
A B C D etc.
  D I R
  D M H
    X R
      Q
      X
```

Or you could make a separate distribution of beginnings (entering the following letter ABOVE the position of the cipher being counted) and ends (just putting a tally BELOW). Or else you may use capital letters for entry when the letter being tallied begins a word and small letters for other digraphs.

In any case, the information obtained may prove very helpful, even showing you which SECOND letters appear most often, for this position (like first and last) has some odd behavior patterns.

The complete count (using check/tallies above final letters) is:

```
    -                                                   
    -       -   -   -   -   -   -                   -   
    -           -                   -                   
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
  D I R L Q   H Q I I W E R R   Y V C   Q M X V Q S
  D M H M Q   W M R   D N M     C U Q   X J Q Q
    X R         Q W       W       O M Q     X F S
      Q         J K       S       R C Q       G F
      X         I U       Y       S V O       X W
                                  B C           I
                                  K             Q
                                  R
                                  S
                                  C
```

Q would certainly be a good bet for E, being the highest frequency letter, and having only one beginning, and a couple of endings. C might be a tentative S (although I would be a fair possibility for plain S).

How about the peculiarly-behaving plain Q? There may be no occurrence of it, and therefore cipher A, P or T could be plain Q; or it could be B or F, or the single Z. If it is B then the cipher D is plain U; if it is F, then cipher Q is U — but we just said it was probably E. So let's keep in mind that B could be plain Q. (You can always use a VERY light pencil when you fill in such possibilities.)

What about cipher H, which has only two occurrences, but as the HH digraph? It could be a rather rare letter, which is often doubled —not J, K, Q, V, W, X or Y — but perhaps Z? (Could be a B, or even a G.)

A mixed sequence sliding against itself has quite different properties from the previously mentioned examples, so there is not much point at this early stage in playing around with alphabet reconstruction. We simply have to get a little more plain text placed. If you have a good computer program, it would take the frequency of cipher letters and order them in plain text frequency in an attempt to match up the best "statistical" plain.

It would have picked the Q for E, the cipher R for plain T, and so on. (The program would probably have a simple instruction for switching assumptions — and tentative R's to be interchanged with the tentative N's. for example.) But since we are dealing with the problem as a pencil-pushing party, let's put in our good guesses (bold) and our "maybe's" (faint).

When we fill in the second word (here in big and little letters to show the strength of our guess) it comes out -uzz-Es. Looks pretty darn good for "puzzles", doesn't it? So let's go.

```
DRVMELWXVJIQY LDHHWQC, DRUQORMNRCI IM JRVXQRI CMWFQSC,
U    PL    E  PUZZLES  U  E     S         E   S L   S

ZSQBDQRIWG SQBDXSQ QKIQRCXFQ NMSO, IJK EMYQSR JUXWXIXQC.
  EQUE      EQU  E E  E S  E              E      L   ES
```

Don't you agree that the sixth word from the end is probably REQUIRE? You'll note that this is a "pattern" word wherein the first and sixth are the same and the second and seventh are also the same. This is a very powerful entry point.

Again, if you were fortunate enough to have the right computer program, or alternatively a list of pattern words, sorted by length and pattern, you might have tried REQUIRE among a few quick tests.

```
DRVMELWXVJIQY LDHHWQC, DRUQORMNRCI IM JRVXQRI CMWFQSC,
U    PLI   E  PUZZLES  U  E     S        IE   S L ERS

ZSQBDQRIWG SQBDXSQ QKIQRCXFQ NMSO, IJK EMYQSR JUXWXIXQC.
 REQUE  L  REQUIRE E  E SI E   R          ER    ILI IES
```

Look at the word in front of REQUIRE. The frequency of the unsolved cipher letters certainly looks good for the plain of FREQUENTLY, so here we go.

```
DRVMELWXVJIQY LDHHWQC, DRUQORMNRCI IM JRVXQRI CMWFQSC,
UN   PLI  TE  PUZZLES  UN E N  NST T   N  INT S L ERS

ZSQBDQRIWG SQBDXSQ QKIQRCXFQ NMSO, IJK EMYQSR JUXWXIXQC.
FREQUENTLY REQUIRE E TE SI E   R   T      ER    I ITIES
```

I'm sure you see the rest of the letters. The English language is rather redundant in the sense that you can almost "read" text with this many holes in it. (Remember what Hebrew scholars have to do when they read text with no — or at least few — vowel

sounds. You would have no trouble in reading NW S TH TM FR LL GD MN T CM T TH D F TH PRT, which is much harder than the average sentence.)

So we have the text; and at least as a matter of tidiness, if not curiosity, we should recover the alphabet used. Since the Plain and Cipher alphabets are the same, and slid against each other by a certain amount, it is easy to see that all of our cipher-plain combinations are offset by the same amount. Thus, if D is two letters away from U, R is two letters away from N, V is two letters away from C and so on.

THIS IS A VERY IMPORTANT POINT, AND UNDERLIES THE PHENOMENA OF DIRECT AND INDIRECT SYMMETRY WHICH WILL BE DISCUSSED LATER. If the idea is not crystal clear, fiddle around with pencil and paper and prove the point for yourself.

Let's make a chain of the pairs we got from solving the message, in the sense that D goes to U, U goes to B, B goes to Q, and so on. (You will find that you have two long bits of the chain, and a short bit, since there is no cipher occurrence of T, P, or A to enable you to go further, and no plain G, H or J.)

G Y D U B Q E M O K X I T

H Z F V C S R N W L P and JA

What we have is a "decimation" of the alphabet — again, a very important point for you to appreciate. The final alphabet may go Y--D--W, or Y---D---W, or whatever, but we will soon determine which decimation we happened to start with.

I'll be darned. By dumb luck (as happens so often in cryptanalysis) we broke our chains (above) at a place that makes for an eye-catching phenomenon. Notice that we have written G over H, Y over Z, D over F, etc. They are obviously consecutive letters in our alphabet.

So we say BCDFGH(IJ?)KLMNQSUVYZ for the part remaining after the keyword, and know that the combinations ER, OW, and XP and the letter T must be in the keyword. Our original decimation can be corrected to that shown on our partial sequence: G--------Y--------D, so we can write the chains out on a 26-wide cycle, skipping 11 places between the letters of the plain/cipher pairs we established:

G I K M Q U Y T X O E B D

which interleaves with the other parts of the chain to yield:

G H I J K L M N Q S U V Y Z T . X P O W E R B C D F

"A" being placed in the hole, we read TAXPOWER.

Monoalphabetic Substitutions
Using Different Mixed Sequences
For Plain and Cipher (K-4)

Now for the hardest of the four ways of handling alphabets (in the terminology of K1 to K4.) I suppose it is no more difficult than the method used by some of the less professional puzzle magazines, which sometimes assign cipher values at random, or give to the first plain letter that appears, the cipher value of A, the second plain the value of B, and so on.

But since half the fun of solving such cryptograms lies in taking advantage of tricky techniques, and recovering the keys used, we will assume that there is a method behind the madness of any puzzle setter who contributed the messages we are looking at.

Here's one that has no tip, but is titled "Look behind you."

YXAUSX PYYONXFT APFX THOGUVSX GRUC AUSXT NPCWHNGOCZ
THFIXOSSUCNX. JPAUC YPSSPJXF OZCPFXW VL AUCL NFOAOCUST.

The patterned word YPSSPJXF will probably offer some hope of attack, but first let's take the usual frequency count:

```
                                                        -
                                                    -   -
    -                   -                       -   -
    -       -           -                       -       -   -   -
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
U   W       T R N X P       X N F   U X H S L H A X C
P   Z       X O F   X       P C C     X   C     F Y
U   N       I               G S A     S   S     T O
U   P       X               X Z S     U   C     O P
U   L       O               F A J     S   C     F
O   U                         C F     P   C     W
                              T   S
6   8       6 2 2 2 2   2   5 6 6   1 7 4 7 1 2 10 4 2
```

Notice how many repeated digraphs show up? When you examine the text, you have even longer repetitions:

YXAUSX and AUSXT, for example. There's a hint here that T might be plain S, which is borne out by the several "T" endings.

AUCL has a trigraph hit with JPAUC, and APFX hits with OZCPFXW.

Sometimes I write down such combinations as we have just noted, and go through lists looking for a four-letter word, the first three letters of which form the last three of a five-letter word. Unfortunately, no cipher letter stands out as a likely candidate for E, so we'll use other ways to pin things down.

General rules of thumb suggest that C or X might be E or T, but not A, I, or O. (The latter three vowels are not good ending letters.)

There are few two-letter words in which the frequencies are as low as the VL in our problem. The word might be BY, or MY — the chance of either letter being a "main" vowel is not good. If this is so, consider again the JPAUC AUCL we mentioned above. The word would end in plain Y, and there are not many which satisfy us for tying in with the five-letter word.

VERY and COVER/LIVER/MOVER might be acceptable, or MANY and WOMAN/ROMAN. ARMY and CHARM would fit, but the frequencies are not good matches. DENY? GREY would only work if this were a British cryptogram, with --GRE and their spelling of grey. A few others would be possible, but far-fetched.

It would be a big help if we could pin down some of the letters into either the vowel or consonant category. Perhaps now is not the time to digress into a long discussion of this, but at some time the solver must learn a little about "letter cohesion" and such.

Briefly, you can state that some letters have a strong affinity, and some hate each other. RE and ER are both very common. But sometimes a letter will precede a "friend" quite often, but seldom follow it — NT appearing in both short and long words, but TN is not a particularly good digraph.

There are many digraphic charts available which show frequencies established after examination of many thousands of words. If necessary the solver can pursue this line at length, but our sample here probably doesn't justify the effort. Let us merely note the obvious fact that in general vowel-consonant or consonant-vowel relationships are better than vowel-vowel, though many consonant-consonant digraphs are fairly good.

Moreover, out of the 25 vowel-vowel combinations (let's forget W and Y for now) there are only a few that one would be interested in — AI, possibly AU, EA, EI, IA, IE, IO, (great.), OU and a few others. (Naturally many more are possible, though not probable.)

So one can look for certain information, and chart certain possibilities when a pattern such as PSSP is spotted (as we saw above.) Almost certainly either P is a vowel and S is a consonant, or S is vowel and P a consonant. One can mark the very low frequency cipher occurrences, and if a spot is noticed where a word has a rare letter in the first and third, and a high frequency 2nd value, the chances are good that this 2nd letter is a vowel.

If you wrote down the letter frequencies for a few cipher "words" in our message, it would start out very uninterestingly 4 9 6 7 7 9 6 9 6 6 4 4 6 5 9 6 2, etc.

Had it been something like 3 8 2 9 1 3 7 6 we could have said that the second and fourth letters were probably vowels, but the count here is rather flat, and doesn't help much. Better follow the ideas of the repeated trigraph we noted, and the word BY and the letter S.

Also, why not consider both the VER and MAN possibilities for AUC? But if C or X is likely to be E, we have eliminated C on the AUCL supposition (unless JPAUC is a word like LITHE) and therefore MANY is our best bet — in little letters?

And if all this is true, why not say the fourth word ends with BLE?

```
YXAUSX PYYONXFT APFX THOGUVSX GRUC AUSXT NPCWHNGOCZ
 EmaLE         S m  E S   aBLE   an maLES   n     n

THFIXOSSUCNX. JPAUC YPSSPJXF OZCPFXW VL AUCL NFOAOCUST.
S   E LLan E    man   LL       n     BY manY    m naLS
```

Looks great, doesn't it? I'm sure you will have no trouble in getting the rest of the text out, but how about the alphabets?

Remember they will have different keywords, but in each one a good deal of the normal sequence might be vaguely apparent, with a few misses, such as a string like DFGJKLMPQ. I find that if I write a normal alphabet down, and put the plain/cipher (or cipher/plain) pairs reading down to the normal alphabet but repeat them from the normal alphabet downwards again, I may see something.

This sounds confusing, but if you put Y over the normal alphabet F, then find Y on the normal alphabet and put an F right underneath it, you'll see what I mean. By this method we get:

```
U V N W X Y Z R O     S A C P     F T G H I J   L
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
M   N     R T U V W   Y   C I O   H L S A B D E F G
```

Obviously C is in the bottom keyword, since the UVWXYZ on top do not involve that letter. S on the bottom sequence is similarly in the keyword.

Trying to put the best arrangement down doesn't quite show us the keywords.

```
A?C d?e?F G H I J K L m?N P?Q T U V W X Y Z R?O?     S?
M?N?p?Q R T U V W x Y z?C O?? S A B D E F G H?I? j k L?
```

Q is almost certainly not in the bottom key, although QU could be in the top key if A is in the bottom key.

In the light of the text, could the keys be COPS and ROBBERS?

```
R O B E S A C D F G H I J K L M N P Q T U V W X Y Z
H I J K L M N Q R T U V W X Y Z C O P S A B D E F G
```

Just to clinch all the things you have naturally easily assimilated, we should look at a mono which is not labeled as being K1 to K4, but is presumably one of them. We might encounter a case of a normal alphabet being slid against a reverse normal, but this is almost as simple as a Caesar slide, if you think about it.

If you suspect a normal-against-reverse encipherment, all you have to do is set A to Z against Z to A, and "decipher" the first word as a tentative step.

Then extend the pseudo-cipher you have just recovered (by a normal sequence), and look for some line on the extensions that seems to be plain text.

Suppose the first enciphered word of a cryptogram is DPS. Refer to the following juxtaposition, and "decipher" the letters:

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Z Y X W V U T S R Q P O N M L K J I H G F E D C B A
```

DPS is turned into WKH by this procedure, and the next step is to extend these letters on a normal alphabet. For economy of space I will extend to the right rather than down:

```
W X Y Z A B C D E F G H I J K L M N O P Q R S T
K L M N O P Q R S T U V W X Y Z A B C D E F G H
H I J K L M N O P Q R S T U V W X Y Z A B C D E
```

I stopped when "THE" showed up. Occasionally on short words you may have a choice — the message might be talking about ESP, or perhaps a gentleman by the name of MAX is the subject of discussion.

The setting of the two alphabets proves to be A over W:

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
W V U T S R Q P O N M L K J I H G F E D C B A Z Y X
```

The message we are going to look at has no indication of K1 or such, but does have a tip — in Caesar form, GZKE.

```
BXQ YDC CUDV: CLV ZMUBC LXOZ NZ NDU OMFVB MB KNWCUNOOVQ YI
NDU TXUVWCB; CLV BVKNWQ LXOZ YI NDU KLMOQUVW.
```

Two important things are quickly noted: There are repetitions (CLV and LXOZ and NDU) and the last group looks "funny".

The latter statement might better be described as a phenomenon in which a word is replaced by a string of letters which are adjacent in the normal order, with certain "skips".

What does this phenomenon imply? A little thought will argue for a setup of alphabets in which KLMOQUVW is part of the "leftovers" on the cipher sequence, and is set immediately under the 8-letter keyword of the plain sequence. The method of encipherment is therefore not K1 or K2, since neither component is normal — as evidenced by the missing letters and the fact that a message is unlikely to end with a plain text of ABCDEFGH or something similar. We cannot tell at this point whether we have a K3 or a K4 encipherment.

So let's look at the cipher. You have undoubtedly told yourself by now that the tip is HALF, since when the plain is slid such a small position, the "one away" letters almost read themselves to you.

Unfortunately there are two 4-letter words in our message. We could either take a frequency count, which would probably throw out one of the possibilities, or we can try them both at once, looking for usage of the cipher letters in other spots.

We'll choose the latter suggestion. (I'll bet on the LXOZ occurrence, since one of the two locations of this repeat is probably followed by OF, and NDU may be THE, although I would want to look at the frequency chart before deciding anything like that.)

When we try the CUDV = HALF setting, it makes the NDU repeat read -LA, which is hardly a promising trigraph. So putting in the other possible crib (or tip) placement, the OF is confirmed, with OUR as its probable follower, and the repeat of CLV is undoubtedly THE:

```
BXQ YDC CUDV: CLV ZMUBC LXOZ NZ NDU OMFVB MB KNWCUNOOVQ
 A   UT TRUE  THE F   T HALF OF OUR L  E      O TROLLE

 YI NDU TXUVWCB; CLV BVKNWQ LXOZ YI NDU KLMOQUVW.
    OUR  ARE T   THE  E O   HALF    OUR  H L RE
```

Filling in the rest is no chore, and if we write the pairs (in "downward direction" as before) we get:

```
X Y K Q V Z   L M     O   W N T   U B C D F     I.
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
  S T U   V     Y     H I O L         P R E N A B F
```

With the following to go on, you should be able to figure out the two alphabets — the case being K4, rather obviously:

```
v w X Y Z              B C D F G H I
e n A B F              S T U V w x Y z
```

Hint: For the lower alphabet, what keyword has the missing values C, D, E and no value above R? And remember that the keyword alphabets often have something to do with the text in CRYPTOGRAM ciphers.

Foreign Language Problems

A word might be said about foreign language puzzles and ciphers. In the matter of crossword puzzles, even a competent linguist may have trouble with languages other than his native tongue. I may know how to get an idea across when traveling in Germany or France, but when it comes to the fine nuances of the language, or the exact word for the erasing thing that fits on the end of a pencil, I may easily get lost.

Out of the many English synonyms for BASE would you expect a Spaniard to associate the word with SECOND (which has a few meanings on its own) as easily as a native-born American could?

On the other hand, foreign languages don't pose the problems for a cryptanalyst to the extent that most laymen expect. This is especially true in "official" communications. A message number is a message number, no matter how you pronounce the "thing" that identifies it, whether it be a word in Mongolian or a group in the code used by Lower Slobovia's Military Attache.

The mark we put at the end of a sentence may be called "full stop" by the British, or "punkt" by the Germans, but it has the same function as our "period."

So, many a time you'll find the "system" was solved by a person who would have trouble finding the nearest post office in most countries of the world — although there was probably some assistance during certain phases of the work by a competent linguist (who might not be able to solve the messages you just studied, above.)

Don't be afraid of tackling a Portuguese monoalphabet, or a more complex system (which you may be going to hear about after a few pages) in German.

I'll admit the occurrence of a problem in Afrikaans, or Chaucerian English, or Pidgin causes a sharp drop in the number of solvers when such appear in the pages of *The Cryptogram*, but this may be due to timidity rather than lack of potential ability.

Remember that individual letter frequencies may differ somewhat from English when you're thinking in terms of German (very high on E!) or Portuguese (what about those Os!) or some of the Scandinavian languages that have those funny markings over several of the vowels. But in the chapter at the end, designated as your own solving ground, you may have fun with something in Esperanto or Spanish or Dutch.

Just to whet your appetite, let's look at a mono in French. A tip is given — L'ON"; and the subject matter is presumably "FLEURS". We'll forget about the fact that they may not tell you whether it's a, K-1 or K-2, or K-anything.

MO WMVFB R'VCD FSV UFCZEFV VRBZDV QIFB MVC KVFJ.
EFVMEFVC WMVFBC R'VCD TVLO FSV QOBDZDZIS, FS RISRVBD
EFV M'IS TISSV O EFZ M'IS GVFD YISIBVB.

You don't have to know a lot about French. The order of letter frequency is slightly different (the one I use begins EAISTNRULOD) but you could stumble along on your feelings for English frequency and not go too far adrift.

It's a good idea to know the articles (LE, LA and the plural LES — remembering the singular LE and LA drop their second letters and substitute apostrophe, if the next word begins with a vowel.) And of course there is UN and UNE. A few verbs of the "to have" and "to be" type would be nice to learn, but the forms of AVOIR and ETRE start branching out, in comparison with our limited English forms. CE (like LE) drops the vowel when it is logical, as in C'EST (IT IS). That's enough for you to go on.

Looking at the monoalphabetic substitution given, there is only one possibility for the tip. This gives LE or LA for the first word, and since the cipher O appears as a

single-letter word, we opt for the recognizable as opposed to the unknown E. And while we're at it, we might as well fill in the logical LES and UNE, which agree on the cipher V decision.

```
MO WMVFB R'VCD FSV UFCZEFV  VRBZDV QIFB MVC KVFJ. EFVMEFVC
LA  LEU    ES  UNE  U   UE  E    E  OU  LES  EU     EL  ES

WMVFBC R'VCD TVLO FSV QOBDZDZIS, FS RISRVBD EFV M'IS TISSV
 LEU S   ES   E A UNE  A     ON  UN  ON E     E L ON  ONNE

O EFZ M'IS GVFD YISIBVB.
A  U  L ON  EU   ONO E
```

Obviously the title "fleurs" and the "c'est" give you a few more critical letters, and you will come up with a pairing of cipher/ plain which can fit over and under the normal, to give us a look at the alphabet (or alphabets).

```
O   R   V W     Z     M U S I Q E B C D F G   J K
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
  R S T Q U V   Y Z     L   A   P C       E F     I
```

There seems to be no question that this was a K-2 substitution, with the plain being normal, and the keyword coming from A LA MUSIQUE.

The Headline Puzzle

Although I gave Paul Derthick a few paragraphs in Part I, showing his interesting variation on Monoalphabets which he labels "Headline Puzzles," we can expand a bit on the earlier quick going-over. They are still appearing in the house organization of the agency for which we both worked (he and his wife being retired, but both following their avid interests in word puzzles of all kinds.)

The idea consists of taking headlines from a recent daily paper, and enciphering five of them with monoalphabetic substitution, the gimmick being that each headline was enciphered at a different setting of the same mixed sequence (against itself).

Thus, each "message" was a type we have discussed as being K-3 — a mixed sequence sliding against itself. Because some of the headlines are rather short, the task might be fairly difficult if each "message" had to be wrestled with as a separate and unique problem, but the fact that they are related (by means of the common sequence) gives an extra leverage.

The weakest link in the chain of five may be broken first, and then the problem of reading the other four is speeded up almost immeasurably. Although you will meet in a subsequent chapter an encipherment categorized as polyalphabetic substitution wherein five or more settings of the same mixed sequence are used one after another, the process is rather more complicated than the Headline Puzzle we are presently

discussing; yet it has certain elements which smack of the solution we are about to discuss. Perhaps the Headline is a problem placing itself half way between the monoalphabetic group we have just finished, and the polyalphabetic category we are about to meet a few pages further on. Let's examine a set of five he recently sent me:

1. EQGDIFPP WKCGP MQ TGNFPMTDCMF WQNFIMA WQKTETFP
2. ZIXGH XUNCZ ERNQI EJU ZIYICZ' URNQIZ
3. MAOGYVM VYCK, MZRI AR TGNMARG, VYG OGARI YGUAUGT
4. LTKDS HYXPJHKJE GJBJMKL HGD STEPTE YJTMJ XOOJG
5. GFOGTFED XGOZPT XFEOO KEMVFETM KPC XACIEIGO

Obviously we would do well to cast an inquiring glance at each of the five in turn, looking for some weak spot. Perhaps we will find an obvious "possessive" (with an apostrophe before the last letter) or a few short words that have letters in common, such as TO, OF and FOR.

Remembering that we are dealing with current headlines from various pages of the paper — including sports, politics, national and world events — we may spot patterns for such as YANKEES or DODGERS or STRAWBERRY or SALVADOR and NICARAGUA or whatever. In headlines, as in short sentences, there is a certain rhythm to the grammar, with a noun-verb relationship that often shows someone or something is about to do something to someone or something. You would be amazed at the number of times the second word ends with S and the third is a short word such as TO .

In the first one given above, you will note the the double letter ending of the first word shows up as the single end of the second word. What better place for an S, which also shows up as the last letter of the headline. And that two-letter word could easily be TO. Let's try it.

```
EQGDIFPP WKCGP MQ TGNFPMTDCMF WQNFIMA WQKTETFP
 O    SS     S TO      T   T   O   T   O     S
```

Note the frequency of F would make it a nice candidate for E (to fit before the SS):

```
EQGDIFPP WKCGP MQ TGNFPMTDCMF WQNFIMA WQKTETFP
 O   ESS     S TO    E T   TE  O E T   O    ES
```

Call it a lucky guess, but I tried CONGRESS for the first word and things fell to pieces quickly.

```
EQGDIFPP WKCGP MQ TGNFPMTDCMF WQNFIMA WQKTETFP
CONGRESS plaNS TO iNvESTiGaTE pOvERTy pOLiCiES
```

Now we see how much of a sequence we can reconstruct, by chaining the cipher-plain pairs:

FECAY QO DGNV WPS KL and MTIR

are the result.

The next step involves the very important aspect of SYMMETRY, discussed in Part I, but the value of which cannot be overestimated, and will bear much reiteration.

We can use the small fragments of the partially recovered alphabet in the process of breaking into some other message. One way to do this is to look for a word (in the cipher) which has adjacent (or nearby) letters which appear close to each other in the re-covered fragments.

For instance, the VYG in the third message has V and G which appear in our recovered sequence DGNV.

This means that the plain letters must be spaced the same distance on the final sequence as V and G — we may not have them at present, but we might.

Remember this is true regardless of the fact that there is little likelihood of our having stumbled on the correct decimation of the final sequence. Perhaps instead of FECAY, we should have written F.E.C.A.Y — or F..E..C..A..Y.

There are 26 ways to spread these relationships out, although many of them are practically equivalent. Eventually we will determine which decimation is correct, but for now, all we need say is that the distance between F and E is the same distance as between A and Y, and so on.

Is this spatial relation perfectly clear? Looking at our scrappy bits of recovery, can you see that F to E is the same distance as D to G or T to I? By the same token, A to F is the same as Y to E, or V to D, or R to M.

So we can consider in turn the combinations C-F, A-E, Y-C, N-D, S-W, I-M and R-T for possible three-letter words — taking all the segments and trying any letter where we have the plain value two positions in front of it, which is the distance the V-G appears in DGNV.

Obviously the only one which looks likely is ArE (unless we are talking RoT, or finding a RaT in some political deal).

I find an easy way to relate two sets of short sequences (since we can add the information gained by assuming VYG in the third message to be ARE) is to take the horizontal strings first recovered, and run them down vertically, pairing the new "crib" letters horizontally:

```
d
G E m
n c t
V A i
  Y R
```

If G goes to E, and V to A, the N must go to C; also since Y goes to R in our new relationship, we can add the MTIR sequence originally recovered, and thus E goes to M, C to T and A to I. We will try everything we can in Message 3:

```
MAOGYVM VYCK, MZRI AR TGNMARG, VYG OGARI YGUAUGT
 i E A  ARt        i   EC i E  ARE  Ei   RE i E
```

Everything looks rosy. The -EC-i-E answer shows lots of promise. How about DECLINE? (It can't be DECEIVE, since the second E would not agree — obviously the pattern is not correct.)

```
MAOGYVM VYCK, MZRI AR TGNMARG, VYG OGARI YGUAUGT
LI ERAL ART   L N  IN DECLINE  ARE  EIN  RE I ED
```

The final plain stands out like a sort thumb:

```
MAOGYVM VYCK, MZRI AR TGNMARG, VYG OGARI YGUAUGT
LIbERAL ARts  LoNg IN DECLINE  ARE bEINg REvIvED
```

At this point, it might be wise to see how the tableau of plain against the five cipher messages is beginning to look. Since we have only fragments of the sequence, we have to write rather sketchy boxes:

	Plain: MTIRU / WPS.FECAY / KLDGNV / QO / B				
1	MTIRU	WPS.FECAY	KLDGNV	QO	B
2					
3	WPS.FECAY	KLDGNV	MTIRU	Z	QO

I haven't tried to fit these bits and pieces into logical order on a final 26-wide sequence yet. At this point we aren't sure whether or not we have bits of one long chain, or of two 13-wide segments — which depends on whether we have an odd or an even decimation of the real alphabet. (We know it is not a case of a decimation of 13, which breaks down into 13 reciprocal pairs however.)

So we had better tackle some other message, and the most likely one seems to be #2. The reason for this is the apostrophe of ZIYICZ' —such appearance often indicates the letter S, as you may remember from certain rules of usage concerning the possessive case. Let's put the guess in:

```
ZIXGH XUNCZ ERNQI EJU ZIYICZ' URNQIZ
S       S           S   S       S
```

This looks good for several reasons — several plural endings, for instance. Notice the occurrences of ZI and IZ; this would do well for an ST combination (or conceivably an SE pair). Let's try the ST.

But looking at our skeletonic tableau, we see that if I is T, we get a flock of other values — for I follows T on our sequence, which means we can fill in all adjacencies:

	Plain MTIRU / WPS.FECAY / KLDGNV / QO / B				
1	MTIRU	WPS.FECAY	KLDGNV	QO	B
2	MTIRU	WPS.FECAY	KLDGNV	QO	B
3	WPS.FECAY	KLDGNV	MTIRU	Z	QO

```
ZIXGH XUNCZ ERNQI EJU ZIYICZ' URNQIZ
ST D  RGES FIG T F R STATES  RIG TS
```

The plain text is obvious, and the alphabets now look like:

	W P S Z F E C A Y H Q O J M T I R U X B K L D G N V
1	V W P S Z F E C A Y H Q O J M T I R U X B K L D G N
2	P S Z F E C A Y H Q O J M T I R U X B K L D G N V W
3	X B K L D G N V W P S Z F E C A Y H Q O J M T I R U

Now, reading the last two messages should be a cinch, since all we have to do is "complete the plain component" of some word. Take the first word of Message 4, for example, LTKDS. If you extend the cipher on the sequence you have derived, you will come to a plain you can use to get the setting. (LTKDS becomes DILGZ, then GRDNF, and so forth.) You might as well have the fun of setting the last two messages.

Without going into great detail at this moment, I might mention the fact that Paul uses a 5-letter keyword for setting the alphabets, but sometimes tricks you with a reversal or something. Also, the main alphabet is mixed up in a tricky way, with a transposition box — but this is not the place to dwell on alphabet reconstruction, as you will find it covered in a separate chapter. I will append a fresh Headline Puzzle in "Problems to Solve," since this innovation is good fun and good practice.

NOTES ON CHAPER III

(1) You might be surprised to learn that occasionally a very profound knowledge of a foreign language almost serves as a handicap rather than a big help, in certain tricky situations.

Naturally, I speak in half jest, since we were so often indebted to the splendid linguists who worked side by side with us.

But there were times when we had to remind our scholarly specialists that the poor guys pounding out the messages we were examining were possibly non-college graduates, and might not be quite up on the subtleties of fourth declension verbs, for example.

I remember Dr. Pip, as we called our one-and-only translator in the section I first served in, clicking his tongue at some of my more outrageous stabs at solution, and saying "but they don't say it that way."

Who don't say what what way?

When I pieced together a keyword that looked reasonable — the CH's and the LK's all behaving as I thought — my mentor shook his head, since there was no such word. And the two halves made no sense when put together. Only when I managed to find the combination in the fine print-footnote of a very large dictionary could I satisfy the "authority" — the word was an obscure machine used in the hat-felting industry.

Since my knowledge of German was mainly derived from such as "Ist Das Nicht Ein Wagenrot" and I consistently butcher my French verbs, and I think the only Japanese character I can really draw is "big" because it is a "man" with his arms outstretched, you can appreciate the fact that the linguistic fraternity has given me up for lost.

But I have no trouble explaining to myself that this bad grammar I am trying to foist on the translator is due to the fact that the code clerk, while following instructions as to the general context of the message to be sent, was probably a plowboy with elementary level education, and didn't care all that much about the subjunctive case.

Chapter IV

MONOALPHABETIC SUBSTITUTIONS
(Without Word Division)

WHAT a whale of a difference a word makes. We have just breezed through a number of monoalphabets which had breaks between words, giving you a chance to consider beginning and ending letters, spot phrases such as "of the" and perhaps decide between "I" and "a." From this relatively abundant hunting ground, we are going to jump into an area in which targets are few and your ammunition is limited. However, the game is still approachable, and you should be able to come home with a few trophies in your tucker bag.

These so-called Patristocrats are ciphers on a much higher level than the Aristocrats we have seen. Apart from the the word separation factor, the rules are identical — tips sometimes being given, keys of type K1 to K4 being utilized, and so forth.

But solvers who assume that a mono is a mono, and this second category is probably just a little tougher than the first, are due for a shock.

"Patristocrats" fall into that class of cipher wherein solution may be quick and painless if you have a crib (or at least have a good idea about the subject matter); and without such help, success may be interminably delayed or even close to impossible.

PLAYFAIR in particular, and many other fairly low-level ciphers as well as some commercial cipher machines, have this "weakness." Given a crib, solution may be a matter of minutes. No crib, change that last word to hours — or days, and the "weakness" doesn't impress us so much.

Of course, "in the real world" there are computers and other high-speed tools which would demolish any pretension of difficulty for the examples cited. But we are supposed to be having fun with word-problems which can be solved without recourse to such.

Let's see what we can do with a Patristocrat when a tip is given—GRANDMA in this case:

BILMN OGOIN ZFAIC SYJKN HJXGM ZIJPM ZPETM NMKST IYSVJ OOIOZ
TIKZT IWFFY FVNVJ NEBFG VNUGO ZTJAI ZFNYM KLRYF HJZIJ QGW

You will note that 5-letter grouping is standard procedure, and has nothing to do with word length.[1]

Rather than our taking the complicated "digraph" count illustrated in previous monoalphabets, where endings are particularly important, let us merely take a quick check-and-tally, or better still, take advantage of the fact that a count has been provided for us:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
2	2	1		3	7	5	2	10	8	4	2	5	8	6	2	1	1	3	5	1	4	2	1	5	8

For certain purposes I prefer the type of count listed above rather than a simple listing of the letters by frequency order:

```
10= I     8= J N Z     7= F     6= O etc.
```

One may gain an overall view (such as the fact that this particular count is almost TOO flat, and suggests a type of text similar to "THE QUICK BROWN FOX" wherein an attempt is made to use all the letters of the alphabet.) Or occasionally one may spot a hint that one of the components is normal, because of the bulging of values where the keyword begins, preceded by the paucity of values one expects at the tail end of the alphabet.

Fortunately, the tip has a pattern of sorts, though not one which is likely to have a unique placement. So we will examine the spots in which we have a repeated letter with three intervening letters. Actually, the crib can be "patterned" as being ABCDEFC — unique letters except for the repeat of the third letter as the seventh.

Let's look at the cipher written without the meaningless blanks (since they may have served their purpose in helping us write a correct text.)

```
BILMNOGOINZFAICSYJKNHJXGMZIJPMZPETMNMKSTIYSVJOOIOZ

TIKZTIWFFYFVNVJNEBFGVNUGOZTJAIZFNYMKLRYFHJZIJQGW
```

You will note I did not bother to mark the I—I at the carry-over between lines, since the repeated "O" immediately preceding would not be compatible with the GR plain.Similarly the placement at IOZTIKZ or OZTIKZT would not fit the crib.

Looking at the frequency of the letters indicated as possible placements, one would hope the best position might fit the LHHHMLH generalization of low, medium, and high expectancy of plain values.

```
SYJKNHJ 3-5-8-4-8-2-8     MKSTIYS 5-5-3-5-10-5-3 and
BFGVNUG 2-7-5-4-8-1-5
```

The first and third of these seem reasonable, but the second does not fit well at all.Let's put both "good" guesses in, and while we're at it, try the best cipher value (I) for the letter E (in small letters) and the only other likely letter not used (z) as a possible E also:

```
BILMNOGOINZFAICSYJKNHJXGMZIJPMZPETMNMKSTIYSVJOOIOZ
 e  D   eD   e GRANDMA     e          D NG eR  A  e
G   D A  DeR        D   A e    e     D              e

TIKZTIWFFYFVNVJNEBFGVNUGOZTJAIZFNYMKLRYFHJZIJQGW
 eN  e   R  D AD      D        e  D  N  R    e
   e   RR R D  D GRANDMA e     eRD      R  e   A
```

I like the ED and EN and ER combinations of the first assumption better than most results we get in the second, and the NG is perfectly placed.

Considering the second possibility, the RR-R-D grouping would only interest us if there were an E in at least one of the two holes. (Remember the small letters could fit in either — or neither — alphabets.)

Let's see if any of the assumptions give something that looks good in the way of alphabets:

```
J     N i   S           H K       Y
G     N z   B           U V       F
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
              M e A N     D         G           R
  G       R A             D             M N       e
```

What would you choose out of this mess? As you will recall, if the method of encipherment were K2, you might find a hint of a keyword-mixed alphabet ABOVE the normal as we have written it (cipher sequences mixed, normal plain).

If K1 were the underlying method, you would spot evidence in the letters UNDER the normal. (If K3 or K4 were used, you might or might not see a lot of evidence, bearing in mind K3 would also have no letter "going to itself".

In turn: The top line doesn't show much promise except for the HK over the MN — which in itself would not be too good, since J would have to be in the keyword. Next line looks good for the UV pair, but the B preceding it would have to imply a very long keyword and a short "left-over" stretch.

The third possibility looks rather good if the MEAN..D were in the keyword.The last possibility is not so hot.Agreed?

Shall we concentrate on the third possibility, which implies a K-1 encipherment? (And fill in the M = I implication, too.)

```
C:      A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
P:                    M E A N   I D         G           R

    BILMNOGOINZFAICSYJKNHJXGMZIJPMZPETMNMKSTIYSVJOOIOZ
     E ID   ED   E GRANDMA  I EA I    IDING ERG A  E

    TIKZTIWFFYFVNVJNEBFGVNUGOZTJAIZFNYMKLRYFHJZIJQGW
     EN  E   R  D AD     D     A E  DRIN  R MA EA
```

Notice that -IDING-ER involves the letter T, which should be plain H, either by the GH evidence or the HIDING HER evidence. And with the SVJOOIO deciphering to G-A--E-, you should have no trouble finishing this.

We noted that the problem just given had a very flat count, because the composer tried to use words which would involve X, Q, and so on. Sometimes this boomerangs, since the use of the unusual letters implies an associated common letter quite

frequently, which is often a vowel. For example, considering the letters with frequency 1 or 2 in this mono, there are 16 occurrences in which we can count the letter preceding; out of the 16, there are 4 E's, 3A's, 2 O's, and a U and a Y, with only 5 consonants.

This is naturally more useful in Aristocrats, where the word length is given. In Patristocrats, two vowels can more easily come together in unusual combinations by crossing words, as in "TO EACH ONE OF YOU I AM ..."

Tough as it is, shall we try another PATRISTOCRAT — one with a more reasonable letter frequency, but no "tip"?

```
ONTJKGMOABCEJKMNFTGTQGBLANOKXQSPEWKGPCPWWBXFNBMWLAACTEO
NTWBLOBJUPYPXONTSFTGTRBXMKETGTEOBBYGTRKBLMOBQTTIYBGOTE
```

I ignored the breaks of 5-letter groups, and will give you the letter count (although in practice you would be better off taking a running digraph frequency count — N under O, T under N etc.).

```
A B  C D E F G H I J K L M N O P Q R S T  U V W X Y Z
4 12 3   6 3 8   1 3 6 4 5 6 9 5 3 3 2 14 1   5 4 3
```

You will note the three repetitions of ONT, and will agree that the frequency looks very good for THE:

```
ONTJKGMOABCEJKMNFTGTQGBLANOKXQSPEWKGPCPWWBXFNBMWLAACTEO
THE    T       H E E     HT                  H       E T

NTWBLOBJUPYPXONTSFTGTRBXMKETGTEOBBYGTRKBLMOBQTTIYBGOTE
HE   T       THE  E E          T    E     T  EE    TE
```

Now we have to take some flyers; there's nothing wrong in "guessing" a few letters, as long as you know where you stand — using erasable pencil for "hunches", and perhaps even colors — red for "solid" values, green for "good" for example. Or if you have plenty of room on your cross-section paper, you can have various levels of plain underneath the cipher.

For instance, the last letter looks about right for D or S — perhaps a little low for R (although it could be.). Let's have two levels of text for now, and while we're at it, try a "g" (in small letters) for that cipher A, since GHT is usually what you get with a .HT, and the frequency looks good.

```
ONTJKGMOABCEJKMNFTGTQGBLANOKXQSPEWKGPCPWWBXFNBMWLAACTEO
THE    Tg  d   H E E    gHT     d            H    gg EdT
           s                   s                      s

NTWBLOBJUPYPXONTSFTGTRBXMKETGTEOBBYGTRKBLMOBQTTIYBGOTE
HE   T       THE  E E     dE EdT    E     T  EE    TEd
                          s   s                     s
```

Note the ...gg-Ed followed by THE. The low frequency letter between g and E ought to be L — it is probably too low for I or R, the other and more remote possibilities. And cipher L has to be a vowel — a low frequency vowel perhaps; let's examine the text before trying all four remaining possibilities. This vowel also fits before the gHT we assumed.

IGHT and UGHT are very good, but OGHt and AGHT are lousy. Also, cipher B must be a vowel from TG-L.So, keeping an open mind about cipher B, E, and L, but firming up everything else — but HOLD IT.

There is a double B towards the end, and the only vowel which looks good there is O. And that means that L is plain U. (Notice how quickly things snowball, and you have to scramble to keep up with assumptions that pile up on assumptions.)

```
ONTJKGMOABCEJKMNFTGTQGBLANOKXQSPEWKGPCPWWBXFNBMWLAACTEO
THE     TGOLd    H E E  OUGHT      d     L    O  HO  UGGLEdT
             s                      s                      s

NTWBLOBJUPYPXONTSFTGTRBXMKETGTEOBBYGTRKBLMOBQTTIYBGOTE
HE OUTO       THE   E E O    dE EdTOO   E   OU TO EE   O TEd
                            s    s
```

Now, with GOLD settling the cipher E problem, the OUT OF possibility followed by a noun (place name?) that goes UPYPX with the P almost certainly being a vowel —what more do you need to finish this. Assuming you have read the text, you will have the following cipher-plain pairs (reading downward to and from the normal, as I suggested much earlier — in case the key is K1 or K2.)

```
P Q R E T J A N K U   C W X B Y   G M O L   F I S
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
G O L   D W R   X F I U S H T A B C Y E J   M N P
```

Even though it cannot be K1 or K2, the keyword stares you right in the face. But just to emphasize the technique, let's chain the pairs:

P A G R C L U J F W M S Y P — hey, we're back to the first letter of the chain, P. How many letters have we? 13? Ah, that's the answer.

There are three types of decimation that a sequence 26 long can be subjected to. One is the unique situation wherein the first letter is paired with the 14th (skipping the intervening twelve) which is tantamount to saying "count the letter 13 AWAY".

If you count from the 14th, and skip 12 letters, you naturally end up with the first again. So instead of each letter leading to a different letter which leads to a different letter, you get 13 pairs — on a normal it would be AN BO CP, etc.

Cipher-plain pairs such as this could happen only when the same alphabet is slid against itself at the "half-way" interval, in K3 encipherment.

But we have only 13 letters, then swing back to the first. This is symptomatic of an EVEN decimation — 2 away, 4 away, etc. (Or saying it a different way, one position

skipped — for instance positions 1,3,5,7 etc., which returns to 1 after 25. There are thirteen such decimations on a 26-wide sequence.

If we have an ODD decimation, 1 away, or 3 away, 5 away, etc., you will get a complete 26-wide chain twelve times, and on the "13 away" case you will have the pairs originally mentioned.

So what we found was a K3 usage, at an even decimation. How to make one sensible alphabet out of the two halves? Consider them:

P A G R C L U J F W M S Y and the fragments Q B O T E D and K I X N H, with the two letters V and Z missing. (Perhaps we were lucky to get the chain of 13 doubling back on itself, for with several small chain segments, we can't be sure what is going on.)

Find a letter in one chain which is uncommon, and therefore probably not in the Keyword, and which has the sequent "uncommon" letter in the other part of the chain. For example P and Q, or J and K. Juxtapose the sections so that these letters are over each other:

```
P A G R C L U J F W M S Y
Q B O T E D
            K I X N H
```

We had two cracks at this, and both of them look good. If these pairs are adjacent, the mixed alphabet should look as follows:

A B C E F I J K M N P Q R T U v W X Y z as the part "left over", with the pairs GO LD FI and SH fitting together as they stand, to form the keyword.

Monoalphabet Variations

There are other simple systems, monoalphabetic in nature, that one might confront in addition to the types used as examples. Some magazines and newspapers will print "Cryptograms" of a slightly less sophisticated nature.

Sometimes the composer either does not know that there ought to be some rule of generating the alphabet used for substitution, or does not care. As a result, there may be an unrelated jumble of the 26 letters placed under a normal alphabet for his use in encipherment. This, of course, cannot be recovered by the solver except by text recovery.

In other cases, the composer likes the idea of assigning the first plain value to cipher A, the next to cipher B, and so on. The resulting text looks as follows:

```
ABCDEFGHHI J KELHM CGNOBC OGPB NOB QLF EQ DEHPJFK G SBIKECM.
```

(You might take a crack at solving this. It's not always easy, since one cannot play around with relationships of letters in the Keyword sequence.)

Theoretically, it makes no difference how many symbols are used in the cipher "alphabet", and one could make up a sort of code-like message in which each letter of the plain alphabet was given the name of a state, or an element, or a flower.[2]

As long as each plain value had the same unique cipher equivalent, whether it was a letter or a pictograph or whatever, the result could be turned into a monoalphabetic substitution as we know it, merely by replacing UTAH with cipher A, COLORADO with B, and so on.

But in practical terms, a substitution is usually a one-for-one, or a two-for two (and occasionally a two-for-one) replacement of the plain by a cipher value.

Another sometime misconception is that by enciphering a message by simple substitution, and then re-enciphering the result, using an entirely different substitution, you have increased the security. This is no better than using a random mixed alphabet once.

Of the numerous variations on simple substitution, several are historically well-known. The pig-pen type and others related to it use some sort of matrixes, with the letters of the alphabet inscribed in some order as:

F	N	E
R	K	W
A	L	I

S	D	J
B	G	M
C	H	O

P	U	Y
Q	V	Z
T	X	•

To represent the letter R one simply draws a symbol ⊐ which shows that the letter is on the left hand side of the first box in the middle row. Use of the other boxes entails a dot being placed in the properly shaped symbol if the letter is in the second box, and two dots are placed in the symbol for the third box.

Thus ⊏ is the letter E, but if that L-shaped symbol had one dot in it, the answer would be J, with two dots being used to indicate Y. The simpler versions of ancient times involved fewer letters in normal order.

A cipher called "Checkerboard" uses two letters of cipher to represent one letter of plain by inscribing a keyword in a square and assigning letter names for the rows and columns:

	C	O	D	E	S
B	C	R	Y	P	T
L	A	N	L	S	I
A	B	D	E	F	G
N	H	K	M	O	Q
K	U	V	W	X	Z

CIPHER would come out as BC LS BE NC AD BO. (In most systems involving a 5 x 5 square, J is omitted, and replaced with the letter I when necessary, as in IOHN

IONES.) But you'll agree there is little difference between this and a monoalphabet, since E is always AD, and so forth. (Of course it is less economical, too, if you don't want to spend money on twice as many cipher letters as you have in the plain.)

This can be made a trifle more difficult by using the letters in either order for rows and columns — making DA stand for E, as well as AD — provided that the "name" letters are unique.

Or you can use 10 or even 13 letters for rows and a different set (or the remaining 13) for columns, which might hold the solver up a bit.

When you use more than one letter to identify a row or column, you are getting into an area that is questionably "monoalphabetic with variations".

You could even place a complete square on the left hand side of the "Plain" square, and put another and different square on top, and thus have a choice of five letters for any row and five letters for any column. I don't know of any name for this, but it would be a Checkerboard with variant identifiers, I suppose, and might prove an interesting challenge.

But this would certainly belong in a part of the discussion of "Other Ciphers" including polyalphabetic or some mixed type.

NOTES ON CHAPTER IV

(1) One of the advantages of such regular groupings is that (In addition the the main reason, which relates to communications procedure) the solver also can easily keep track of things. The rather onerous task of copying out groups, counting positions and intervals and cycles, etc. is made a trifle easier by knowing that there are 11 five-letter groups on each line, or some such figure.

Five letters are just about the right-size bite for the transcriber to assimilate, and the analyst gets rather good at treating such a package as an easily-remembered entity.

This is a good spot to give a short warning lecture on the importance of accuracy and consistency, with built-in checks to make sure one has not dropped a letter, or left out a group, etc.

A small error, transposing a couple of letters, or misreading O for Q, is bad enough, but leaving out a group, or dropping a stitch in some other way may be disastrous.

Probably successful cryppies are, on the whole, much less prone to making "stupid mistakes" in their tax returns and bank reconcillations and similar clerical chores than the average person who has not learned the hard way that "it pays to be careful."

(2) We haven't pressed the point, but we trust the present reader has a fair appreciation of the difference between a code and a cipher, even though you will see loose statements in articles on cipher machines, Sherlock Holmes' "Dancing Men," and World War One systems that are misleading. (You will remember that these little "stick figures" Conan Doyle dreamed up resembled miniature Boy Scouts frantically waving semaphore flags.)

In general, the distinction is easy to bear in mind. If a single letter of plain is replaced by a single letter of cipher WHETHER OR NOT IT IS ALWAYS REPLACED BY THE SAME LETTER, it is a substitution CIPHER of some sort.

On the other hand, a code may consist of groups of letters or numbers that "stand for" anything from a letter, a number, a syllable, a word or a punctuation mark to a long explanation of what to do if war breaks out.

Breaking a code may be just too complicated a job for inclusion in the present writing, but the reader may gain an insight into the mechanics (and also the excitement) of the process by referring to Barbara Tuchman's excellent book, *The Zimmermann Telegram*.

This infamous message was sent in the German diplomatic code, just before the U.S. entry into WW-I, and may be the straw that broke the back of resistance to our involvement. Briefly, the Germans extended a carrot

to the Mexican government in the form of promises of U.S. territory if they would cooperate in a successful joint venture against Yankee imperialism.

The British cryptanalytic successes taking place at Room O.B. 40 of the Admiralty had revealed the startling contents of the Zimmermann telegram, and they managed to figure out a way of passing the info on to the Americans, correctly anticipating that it would raise enough dander to effect a positive step towards U.S. involvement.

Returning to the subject of "ciphers," there naturally must be some prearranged way of performing the substitution (or transposition) so that the authorized recipient can reverse the process and turn his cipher message into plain text. Hence, a complicated machine that goes through X number of rotors to perform the task still changes a letter to a letter, and constitutes a CIPHER transformation — which is not a code.

There is a class of CIPHERS which substitute the single plain value by other than a one-for-one replacement. A digraph may replace the single plain, or a symbol of some sort (even a picture of a Semaphore position or the dot-and-dash equivalent.)

We will discuss Polyalphabetic CIPHERS soon, which have more than one way of enciphering any given plain value. And we will talk about ways of changing a digraph to a different digraph, as a substitution process.

In all of the above, the original order of the plain values is preserved, and the clerk at the receiving end "unbuttons" his message without having to juggle anything by placing the first letter last, the second letter first, etc.

But there is a class of CIPHERS in which the single-letter plain value is preserved, but the positions of the various letters are switched around — in small sections (as within groups) or in huge chunks where 144 letters at a time, or more, are scattered hither and yon in the cipher text version.

Naturally there has to be a method in such madness, with prearranged matrixes, transposition keywords, grilles, and the like.

You can even TRANSPOSE an ENCIPHERMENT. But that is not a code.

Again, how do we define a code? Usually a code consists of a specific replacement of a plain value (which may be a letter, a number, a word, a long phrase, punctuation, etc.) by a CODE value. This is usually (but not always) a consistent-length value — 3 letters, or 5 digits, or whatever the system requires.

You can only have a limited number of values for a 3-digit CODE, but a 5-letter CODE has enough room to accommodate anything you might wish to say.

Thus the code group ABASE might stand for the letter T, while the group CARDO might take two pages in the code book, with a general meaning of "War is imminent. Burn your papers. Destroy your communications equipment. Contact a neutral embassy, etc."

Codes can be further enhanced in the name of security by additional treatment, by such means as encipherment (by hand or machine) which turns the product into an ENCIPHERED CODE.

With this brief introduction to the vocabulary of Cryptanalysis, which is usually treated rather loosely, I'm sure you will spot the boo-boo appearing in a well-received book which describes the German ENIGMA as a Transposition System.

I am somewhat taken aback by this, since the book is "Top Secret Ultra" and the author, Peter Calvocoressi, actually worked at Bletchley Park, the British wartime "black chamber." (You can't pick up a book about ENIGMA, or the successful attacks of the Allies on enemy communications, without seeing a picture of the Victorian manor house in Buckinghamshire which sparks so many cryptanalytic memories — mine included.)

I do not know how long he was there, and exactly in what technical capacity he served, but it is possible that I am expecting too much "cryptographic purity" in people who were dragged in from Oxbridge to do very specific work on specific highly technical — but very specialized — projects.

I at first passed the offending statement as a minor slip, or something which happened to get by the proofreaders, but a few pages later he talks about the ENIGMA as a transposition cypher.

I doubt if such as the author ever had the luxury of intensive training in wide areas of general cryptographic diagnosis, or even the history of and practice on "classic" systems. They did what they had to — and we can all thank them for it.

Naturally, when the war was over, the vast majority of the dons went back to their chosen bailiwick — probably with some sigh of relief. And even those with fondest memories of a few years spent in beating the Axis at its own game would probably not classify themselves as long-term professional cryptanalysts.

In other books (and movies) you will also catch references to the "breaking of the spy codes" after the author or script writer has carefully delineated the only system used by the culprits, which happens to be a Transposition.

But am I being too picayunish?

Of course, there is a large class of ciphers which do not change the value of a plain letter — they just change the position of that letter in the message.

These Transposition CIPHERS will be discussed later.

Chapter V

POLYALPHABETIC SUBSTITUTION
Using Normal Cipher and Mixed Plain Components

YOU have proved to yourself, I trust, that the security offered by monoalphabetic substitutions would not be very great in a "real-life" situation.

We have together worked out the solution of several types of such, and you are itching to tackle the samples reserved for you at the end of the book, I am sure. On a higher plane of sophistication and security lies a class of encipherment known as polyalphabetic substitution, wherein several (or many) different alphabets are used to encipher the plaintext.

This can range in complexity from the use of a few alphabets (perhaps even normal ones) being used in a very regular and prescribed way, to a complicated machine utilizing countless alphabets in what may appear to be a random order.

The ones we are about to consider are on the lower end of this scale, using, say, between six and twelve alphabets one after the other, on a regular cycle in some way.

The CRYPTOGRAM calls such problems QUAGMIRES if they have a mixed alphabet in at least one of the components, and borrows the terminology of the keying system we used in monos, above. Thus, we have QUAGMIRE I, which uses a keyword-mixed plain component with a determined number of normal cipher alphabets at different settings; QUAGMIRE II, which uses a normal plain and various settings of the same mixed cipher component; QUAGMIRE III, employing the same mixed alphabet for plain and cipher (juxtaposition repeated on a cycle); and QUAGMIRE IV, which has one mixed alphabet for the plain, and a series of slides of another mixed alphabet for the cipher components.

There are some problems which use normal components on some type of cycle, either direct or reverse alphabets. But the weakness of known components makes them more vulnerable to solution, and they will be discussed in the area dealing with "other methods of encipherment".

Let us take the QUAGMIRES in turn, making sure you understand the method of encipherment as well as the tricks of unraveling the text.

Suppose we build an alphabet on the keyword ENCIPHERMENT:

E N C I P H R M T A B D F G J K L O Q S U V W X Y Z

Let us then take a NORMAL alphabet, and set it so the letter C falls under the first letter of the plain sequence. We will designate this as cipher setting No. 1. Now let us slide the NORMAL to a different position say I under the E of the plain component, for setting No. 2. And so on, in a prearranged fashion — suppose we say in this case there

will be six different slides, with No. 3 starting P, No. 4 starting with H, and numbers 5 and 6 having E and R, respectively, at the first position.

```
Plain:  E N Q I P H R M T A B D F G J K L O Q S U V W X Y Z
Cipher: C D E F G H I J K L M N O P Q R S T U V W X Y Z A B
        I J K L M N O P Q R S T U V W X Y Z A B C D E F G H
        P Q R S T U V W X Y Z A B C D E F G H I J K L M N O
        H I J K L M N O P Q R S T U V W X Y Z A B C D E F G
        E F G H I J K L M N O P Q R S T I V W X U Z A B C D
        R S T U V W X Y Z A B C D E F G H I J K L M N O P Q
```

You of course noticed that the first letters of the six settings (under the first letter of the plain component) spell a word. This is a practice used by communicants who don't want to carry incriminating material around but don't want to overload their memory bank. Remembering "Encipherment" and "Cipher", they have all the equipment necessary to set up business (assuming a pencil and something to write on).

Suppose the first word of the plain message is "REGARDING". The first R is found in the plain sequence , and the letter underneath it in Alphabet #1 is substituted, in this case "I". Similarly, by using the cipher alphabets one after another, returning to No. 1 after using No. 6, the word REGARDING becomes IICQKCFJC. (Note that the plain G's, appearing five letters away, or on a cycle of six, both become the cipher C.)

Now that we know how QUAGMIRE I is made, how do we attack it? Let's look at an example:

```
WBFWX LWVPY WICQJ HJYDL LNABF JCQFB BHMPA XGKIU CRHVK
YNEJO VMDEJ SPQPT GLFFB YOEYD MIHYY JJCPY YDVIE TOFXX
LWPSC YTBKJ ORCYZ DBYDH YHR
```

There is a tip "TLEANDTHENREPLIED" which should appear some place in the text. This is going to be an important entry into the solution, but certain tests should be made to find the cycle (the number of different cipher alphabet settings) regardless of whether you have a crib or not.)

There are at least three ways to determine the cycle — one being simple and the others a trifle more work. The simplest method is applicable if there happens to be a significant repeat in the cipher.

If you spot a 5-letter recurrence of YSNJE, say, at positions numbered 84 and 105 in the text, you can be almost certain that the message uses seven cipher settings, or slides of the cipher alphabet. A little elementary reasoning shows why this is so. The difference between the two numbered positions where the repeat starts is 21. There are only two factors of this number, and 7 is a much more logical cycle for a simple message than 3, which would be far too short for security. Obviously if we wrote the message out on a cycle 7-wide, the repeated text would show up as a 5-letter "hit", each time beginning in the same column of the 7-wide write-out.

Three-letter hits in short cipher messages are probably significant, but not iron-clad. (I will leave it to Dr. Sinkov and other mathematicians to dwell on probabilities, Sigmage of significance, and such. But for our purposes here, if you get a long repeat, be thankful for it, and if you get a lot of short repeats, hope enough of them will be at the correct interval to give you a clue.)

There's not much joy in the repeat department of the message we are studying, since XLW seems the longest one which hits the eye. (I have a computer program which would find all repeats of 3 or more letters, and also perform certain tests to determine the cycle which we will discuss later, but for the time being let's only play with pencil and paper.)

It might be worth your taking a running digraph count, so that you can find all repeats — even digraphs. Two-letter repeats are not significant in themselves, but enough may pile up at a certain interval to interest you.

Let's look at the XLW noted. The first occurrence is at position four, and the second at position 90. This gives a difference of 86, which is not a very encouraging one, factoring to 2 and 43. So much for the guarantee of three-letter repeats.

So let's talk about another method of determining cycle, not so easy as the taking advantage of a long repeat, but not too much trouble.

For the sake of argument, let's assume the cycle is 6, 7 or 8. (It may be 5, or 9 or 10 or even greater.) If we write the crib out on those cycles the three versions are:

awh	awh	awh
ILEAND	ILEANDT	ILEANDTH
THENRE	HENREPL	ENREPLIE
PLIED	IED	D

(You will note that I added a couple of logical letters "in light pencil" — small letters here — to precede the crib. Of course it could be "smile" or something.)

The short crib doesn't generate much in the way of significant repeats to search for. In longer cribs, sometimes several digraphs show up, or even a longer stretch. But we must persevere. If the cycle is 6-wide, the E repeat should show a cipher letter hit with five intervening letters, and perhaps show the validity of AWHILE. The 7-wide test would be easier, since the I over the I and the E, followed in the next line by another E, would give a reasonable pattern of cipher repeats to look for. The 8-wide write-out promises very little, the only help being that there can be NO repeats at the interval 8 within the length of the cycle, which will knock out a few possibilities.

In keeping with our precept "try the simplest thing first," we might look at the 7-wide cycle possibility. In some ways, this sounds like searching under the lamp post for the key dropped someplace along the block — because that's the place where you might see it most easily. But the example we have presented happens to be one of those which has no obvious repeat, in either plain or cipher, and if the "easy test" doesn't work out we will have to bring out the big guns, and prepare for a longer battle.

Our stretch of cipher will show repeats such as ---X----Y----XY--so let's look for the "Y" (not necessarily that letter, of course) repeat with 6 intervening blanks, and then check back on the "X" position, to see if it hits. The easiest way to do this is to follow (either by eye, or with a couple of arrows on a paper strip) the progression of letters, starting say at the second group and looking at the 6th and 13th position of the text, saying to yourself "LC, WQ, VJ, PH" etc.

Or if you are so minded, you can write the text out on two strips, and compare them at the appropriate interval. Some people jump around and take chances, others are more meticulous, and don't mind making a half-hour test if it will prove something that "jitter-bugging" won't.

Regardless of how you do it, you will be happy to note that at one place, at least, you get what you are looking for. (There may be several positions which have to be tested further.)

The stretch "glffbYoeydmihYyjjcpYYdvie" has 13 letters between the Y occurrences which represent I to I, and 6 letters between the Ys which represent E. The coincidental appearance of so many Ys shouldn't throw you.

So we can now write the message out on a cycle which is seven letters wide, and put the plain text underneath the appropriate cipher:

```
WBFWXLW VPYWICQ JHJYDLL NABFJCQ FBBHMPA XGKIUCR HVKYNEJ

OVMDEJS PQPTGLF FBYOEYD MIHYYJJ CPYYDVI ETOFXXL WPSCYTB
              a whILEAN DTHENRE PLIED

KJORCYZ DBYDHYH R
```

Now let's prepare a deciphering tableau, putting the plain values above a normal cipher strip, and using the plain value E to get started. (This choice was a calculated one, since it starts a longer chain of consequences, as you will see.)

```
Plain:              E
C. Alph.#1
         2
         3
         4 U V W X Y Z A B C D E F G H I J K L M N O P Q R S T
         5 A B C D E F G H I J K L M N O P Q R S T U V X X Y Z
         6
         7 F G H I J K L M N O P Q R S T U V W X Y Z A B C D E
```

Since the fourth alphabet also has a plain L, we can enter it on the top line, and similarly place a plain N (from the fifth alphabet). You will note that this N is confirmed by its appearance in the 7th alphabet, so we know we are on the right track.

Since we have the plain L, the second alphabet comes in, too, and hence plain H and T. This gives the third alphabet, and the plain I. Snowballing again?

But there is still more help. We soon get all alphabets except the first, but if we look DOWN the various columns of the matrix, we can easily spot the word COUNTRY, which must have been placed under the first letter of the plain sequence.

```
Plain:    A     D E   H         R   T             L   W I N
C.Alpha1: J K L M N O P Q R S T U V W X Y Z A B C D E F G H I
       2: V W X Y Z A B C D E F G H I J K L M N O P Q R S T U
       3: B C D E F G H I J K L M N O P Q R S T U V W X Y Z A
       4: U V W X Y Z A B C D E F G H I J K L M N O P Q R S T
       5: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
       6: Y Z A B C D E F G H I J K L M N O P Q R S T U V W X
       7: F G H I J K L M N O P Q R S T U V W X Y Z A B C D E
```

We would now be able to decipher any letters we have so far recovered on the alphabets they affect. But a little work on the plain sequence would certainly add the S between R and T, and perhaps help you find either F or G, etc. Rather than belabor the point, I'm going to let you finish the fun.

Polyalphabetic Substitutions Using Normal Plain and Mixed Cipher Components

How about a QUAGMIRE II? This polyalphabet substitution utilizes a Normal Plain and a Keyword-Mixed Cipher alphabet. Let's tackle a problem which gives a tip of TAPHORICORTABOONATUR, and in addition gives the title "Usage":

```
GJGQH JLELW SZGGE TGMQS YVAHU OLFYX NIRJH VKJDS XMZVU EPETG
HIAHW ZOTFN HIHVW QUQDN UENAE QMFQA YXIOV UIVYG NVLUJ MOCVL
RXSOT VSSMT CIIFH VEFYA VJLEU VDQFX OZJHN NUHQY EOGQD YGHEG
RXVVV OBVYY SR
```

There is a slight hint here of the cycle length, in that the crib starts and stops in the middle of a word, and the solver is likely to wonder why.

There happen to be 20 letters given in the tip. Could the puzzlemaker be trying to hint that the width is some factor of 20? Four is too short, five is unlikely when so much is given, since it would be too easy. Ten?

Keep this in mind.

How about repeats in cipher. We note JLE beginning the 6th and the 122nd positions, but this gives a difference of 116, which factors to 58, 29 and nothing that looks good — obviously an accidental repetition. We could take some counts at varying lengths, and this would probably show us the cycle. Is it clear that if the cycle were 6, you could take a distribution of positions 1, 7, 13, 19, 25 and so on, which would have to be statistically rough?

We'll go into the simpler mathematics of this later, but for the time being I'm sure you will realize that a random sample of letters taken out of a hat containing all 26

values (with replacement each time a letter is drawn) would be pretty flat, with as much chance for an X or a Q as there would be for an E.

The same would not be true if you took a count of every fifth letter you found in a newspaper, for instance, or sampled the text in some other way.

There would be about 13% occurrence of the letter E, a few less of T, etc., down to a very few J's and Q's and such. We will later show you how to measure such roughness, or lack of it.

At any rate, the finding of the cycle statistically can be accomplished, and is in fact very easy if you have a computer. But rather than our getting involved in a long check and tally procedure right now, let's see if we can use the crib and the possibly hinted-at width in conjunction with each other.

If we write the crib out 10 wide, we get:

```
TAPHORICOR
TABOONATUR
```

Even without trying to guess what comes before and goes after, this crib has a nice pattern in which the digraphic hit 10 away is followed a few positions later by O over O and the final R over R. That ought to be easy to find.

Sure enough, if you look at the second line and start the crib with the cipher HI, you're in business. Let's look at the message written out 10 wide, and make a tableau in which the plain component is normal, and some mixed sequence is sliding below it with 10 different settings:

```
GJGQHJLELW SZGGETGMQS YVAHUOLFYX NIRJHVKJDS XMZVUEPETG

HIAHWZOTFN HIHVWQUQDN UENAEQMFQA YXIOVUIVYG NVLUJMOCVL
TAPHORICOR TABOONATUR

RXSOTVSSMT CIIFHVEFYA VJLEUVDQFX OZJHNNUHQY EOGQDYGHEG

RXVVVOBVYY SR
```

PLAIN:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C	1																			H						
I	2		I																							
P	3			H												A										
H	4							H							V											
E	5														W											
R	6															Q			Z							
	7	U							O																	
	8			T																Q						
	9														F						D					
	10																	N								

Because we know the plain sequence is normal, and hence must be in the right order, we know that the cipher sequence reads from left to right just as we see it. Therefore the cipher sequence has H----V----A, Q--Z--T, U-----O, and F----D, and we can fill in a few letters on the tableau before deciphering all that is possible.

Anticipating what you would surely do next, I'm adding an N to the IO combination you are going to get, since it also places a V right in front of the W on the sequence we are building up. (I'm sure you would do this on the basis of the two strong bits of evidence.)

If this is true, then the cipher sequence goes either QVW-Z or Q-VWZ. Therefore, in the first position, Q cipher is either Y or Z, and Z cipher is either C or D; in the 3rd line of the cipher tableau Q is G or H and Z is K or L, and so on.

But notice the 6th line in the tableau: Cipher V is either O or P; but V is plain O in the fourth line, and the fourth and sixth setting of the cipher sequences (and thus the Key letter for the setting) would have to be the same.

A quick look at the cipher letters in these two positions shows an interesting "fit" — V, O and Q being repeated in each, with several other letters in common. (There are simple tests to measure the "fitting" of two distributions, but we won't digress here. Rather, let us try it out for size — again being sure we know which is "maybe" and which is "sure".)

But we should also be constantly aware of the text. Note that the plain letter after the crib, – NATUR– can no longer be A (for NATURAL) and the only logical possibility is now E. Slam it in, and try everything at once. (We can always go back to the drawing board.)

```
PLAIN: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
 C  1  V W    Z U      A T         O                 H            Q
 I  2  I
 P  3     H              Q V W    Z U      A T          O
 H  4  O                 H              Q V W    Z U      A T
 E  5                 H               Q V W    Z U      A T          O
 R  6  O                 H              Q V W    Z U      A T
    7  U                    O
    8     A T         O                 H                Q V W    Z U
    9                                     F                  D
   10                                                  N
```

```
GJGQHJLELW SZGGETGMQS YVAHUOLFYX NIRJHVKJDS XMZVUEPETG
   NG            W        PH          A G    U   1O

HIAHWZOTFN HIHVWQUQDN UENAEQMFQA YXIOVUIVYG NVLUJMOCVL
TAPHORICOR TABOONATUR E  V I         ANS U          I
```

```
RXSOTVSSMT CIIFHVEFYA VJLEUVDQFX OZJHNNUHQY EOGQDYGHEG
   AVO       A  GN     A   RO TO  M  H  AH      N   N

RXVVVOBVYY SR
  IONI U
```

Now look at the third alphabet. Where is a good place for cipher G, and what letter (vowel?) would look good in the beginning --NG? Do we both agree on a plain A, putting G before H in the cipher sequence? This gives S-ANG-- in a message titled "Usage" — so how does the cipher J look for plain L in the second alphabet? Can't tell, but I could be in the keyword, and J in the "leftovers." If this is so, the alphabet goes:

```
U--ATI--O-----GHJ----QVW-Z
```

We know that F is in the keyword and D is not (from the fact that D comes a short way after F in the 9th alphabet.) Where does D come? Either just before G, or the sequence is EFG (if E is not in the keyword.) How does D (and perhaps E) behave when we try them tentatively in the skeleton sequence we just derived?

We don't have to go very far to decide that the message begins SLANG IS LANGUAGE WHICH, and that's all I'm going to tell you. If you have been following this by writing the cipher and the tableau on cross-section paper, you will have no trouble completing the problem, getting the cipher sequence, and recovering the keyword used to set the 10 alphabets (probably under plain A, in QUAGMIRE II.)

Polyalphabetic Substitutions Using the Same Mixed Alphabet for Both Plain and Cipher Components

QUAGMIRE III is a very important class of ciphers, because they introduce one of the most powerful tools that Mr. Friedman "invented" and explained in his Riverbank series of papers (which are prized possessions of many amateur cryptographers, in reprint form.) The title associated with this weapon is "Direct and Indirect Symmetry," and we will soon prove its effectiveness.

I explained a little about this in Part I, but the concept is sufficiently important to bear repetition, even if you read that earlier output.

But first we must have something to work on in the way of partial alphabets, obtained by judicious placement of a crib, statistical attacks, and perhaps a few hunches.

The title of the problem is "Inertia in the British labor market", and the cipher given is as follows, with the tip ANDTHREECALLINGFORAMANTOSTANDON:

```
IBWVU PLTPJ TKPPM YCTDV XYGNY QYNTW NFSUI XNACX CFTGV
AIKPS RTCOJ JWPRR VOLAA ARURJ NUIXM XPQBV UIBWO GPCDP
LNNRD FPSLI BUGOC DOTWK CPIRQ RVQGY GCXLV MNOBE QFVOL
GBWGP ATNJL YWRMW EKLAA VICVE AQBKU VFJUR DVIOZ MPTZO
VSLIH QBQXF LLLWH PUSGV XP
```

It's always nice to find a repeat of the first three letters, since it improves the chances of the message starting with "THE". The distance between the occurrences of IBW is 81, which gives no other choice than a 9-wide cycle. (The hit may be accidental, of course, but things look good enough to warrant trying to place the crib on that basis.)

```
A N D T H R E E C
A L L I N G F O R
A M A N T O S T A
N D O N
```

Why don't you write out the cipher nine wide, and look for the tell-tale repeats of the three A's under each other, or 9 letters apart in the cipher. It sometimes pays to write the cipher out on the proper cycle with sequent segments under each other, as:

```
I B W V U P L T P
J T K P P M Y C T
D V X U G M U Q U
```

Remember in QUAGMIRE III, or any case where a cipher component is the same as a plain component, if one cipher-plain pair matches (E is E for example) ALL pairs must match, for the sequence is set A to A, B to B, etc. When this happens, of course we get a column of our write-out as "free plain-text", which is of considerable help.

After you have done a little of your own spade work, we'll pick up the thread of solution, as follows:

```
IBWVUPLTP JTKPPMYCT DVXYGNYQY NTWNFSUIX NACXCFTGV AIKPSRTCO
THE       H  EA  D             E        G T  N   N  E ANDT

JJWPRRVOL AAARURJNU IXMXPQBVU IBWOGPCDP LNNRDFPSL IBUGOCDOT
HREECALLI NGFORAMAN TOSTANDON THE          O    I  H     L

WKCPIRQRV QGYGCXLVM NOBEQFVOL GBWGPATNJ LYWRMWEKL AAVICVEAQ
   E A            O        OLI  HE A NA    EO    I NG

BKUVFJURD VIOZMPTZO VSLIHQBQX FLLLWHPUS GVXP
                N        ND
```

You'll notice I filled in the H of THE, since the other two letters were checked by the crib. The tableau looks as follows:

Plain:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher 1								J						A						I						
2							A	B							X			J								
3					W	A													M							
4					P										R					X						
5	P		R															U								
6	R													Q												
7				B							V	J	T													
8	N			C							O			V												
9									L					U						O						

One cannot overemphasize the importance of our next step. Think back to the reconstruction of the alphabet in the case of K3, when a monoalphabetic substitution involved the sliding of a mixed alphabet against itself. We saw that the plain-cipher pairs had an exact relationship, in that if we had the pairs AD, RG and PH, for example, A to D was the same distance as R to G and P to H on the alphabet strip, or component, or whatever you choose to call it.

The same phenomenon exists here, in that the Plain component and the Cipher 1 alphabet represent pairs that are the same distance removed — H to J, N to A and T to I in this case. Similarly, G to A, H to B, O to X, and R to J are equally separated — though not at the same interval as the first pairs obtained from line 1. (Obviously if H to J is "x" distance, H to B cannot be the same distance.)

Let us write down all the pairs we get by going from plain to cipher in each of the alphabets in turn. We can also write down pairs obtained from "sidewise" relationships For instance A to C on the plain sequence is the same distance as P to R on Row 5. In addition, Row 7 to Row 8 tells us that BC is the same distance apart as VO.

This is the powerful tool in solution of a sequence against itself. You can imagine a little "square" and go up, or down, or across to find relationships within and between both plain and cipher components.

Plain Sequence to Row	1	HJ NA TI
	2	GA HB OX RJ
	3	EW FA SM
	4	EP OR TX
	5	AP CRU
	6	AR NQ
	7	DB LV MJ NT
	8	AN DC LOV
	9	IL NU TO
From Plain A to Plain C		AC PR
From Row 7 to Row 8		BC VO

We have not bothered to list here all the relationships one might obtain. Obviously if we have H to J and N to A, by going from Plain to Row 1, we can get H to N and J to A by going "sideways" on the plain and first cipher sequences.

There is so much information available, one could jump in several directions at once, as far as solution is concerned. In practice, you may find you need every scrap of information you have, and would continue this search for pair relationships, and the combining of such, to the bitter end.

Such eye-catchers as the fact that the relationships of Row 1 in the above listing are the reverse of Row 8 might tempt you to veer off in that direction.

Note that there are several cases of 3-letter chains; if L-O is the same as O-V we can get a three-letter segment, for instance.

Do you see that pairs in the 7th listing above are separated by one letter in a sequence obtained from the next set, as evidenced by LV in 7 and LOV in 8?. Thus we can add the two together as follows:

DCB LOV M-J AN-T

Look at these fragments, and you will realize that by pure dumb luck we have stumbled on a great deal of information about the sequence. Do you see that (in all probability) these fragments are reversed, and the sequence has BCD, VOL, JKM (since we have used L) and T-NA in it?

This robs me of the opportunity to discuss decimations again, and show you how to find the correct one, since it displayed itself so brazenly. (Later on.)

Remember the very important part of the tool of symmetry — that because the plain and all the cipher alphabets are the same, we can associate pairs in any "straight" direction we find them — up and down, sideways, etc. utilizing the plain sequence as well as the nine cipher sequences.

(We will later on see that you cannot utilize the plain sequence in the equivalent operation for polyalphabets of the QUAGMIRE-IV variety, where the plain and the cipher sequences are different.)

Continuing our rapid recovery: We find that A to N along the plain text sequence is the same distance as R to Q in alphabet 6. This means we can add QR to our fragments.

More importantly, we see that the relationship between the plain sequence and alphabet 1 (reading down) involves NA, so we can add HJ and IT to our pairs.

VOL TINA BCD HJKM QR

Notice H to B and G to A on the plain-alphabet 2 relationship? This says that if we (logically) put G just ahead of H, then A goes just ahead of B, as one might expect. (Of course if the keyword has G and A at the right interval we may come a-cropper, but (as we keep on saying), there is always that eraser.).

VOL/TINABCD/GHIJM/QR

where we cannot be sure whether one or two letters have been left out, or perhaps there is none missing.

No, that last statement is incorrect. Since O is in VOL and N in TINA, the only missing element is P. This means that either the part of the sequence which we have goes

VOL/?/TINABCD(f?)GHIJMPQR

or there is no gap

VOL/?/TINABCDGHIMQR

but the first looks so good that we should set up a new (and partial) tableau:

```
      V O L   /?/        T I N A B C D F G H J M P Q R S
1   V O L                T I N A B C D f G H J M P Q R S
2     X                            T I N A B C D F G H J M P Q R S
3                                    T I N A B C D F G H J M
4     Q R S (u and/or W)X
5
6
7
8       V O L                T I N A B C D F G H J M P Q R S
9  ____________________________________________________________
```

— but after starting the thing, I realize it would deprive the reader of the pleasure of carrying out the solution in a straightforward manner. I was going to point out the obvious — that the eighth alphabet starts two places to the right of the first, or one position to the right of the plain, and the relationship of where 1, 2 and 3 start will enable one to put in X for the first and third alphabets, etc. So have fun getting the alphabet (with its keyword) and the second keyword which presumably goes downward under some letter, to show setting.

Polyalphabetic Substitutions Employing a Keyword-Mixed Alphabet for the Plain and a Different Keyword-Mixed Alphabet for the Cipher

This class, known to CRYPTOGRAM readers as QUAGMIRE-IV, is probably the most difficult of the Quagmires, since there are two alphabets to recover (instead of one) and you cannot use the tool of symmetry to cross the boundaries between the two alphabets. (This will become clear as we progress.)

We'll look at a cipher entitled "Lost horsepower" which gives a tip and its location:

Ends: WEENLATESEPTEMBERANDDECEMBEROFTHATYEAR.

I have written the cipher out in 5-letter groups, but remember that we do NOT know the correct cycle — which could logically be anything between 5 and 10.

Because this class is difficult, and you may need a little extra help on your first assignment, we'll give an extra tip: It begins THEGREAT...

```
MWQYD KMCAO KHSEE YULIH WYTEW YRLHG LMEJC ZHAKE NYWUP
THEGR EAT

QSQSO ESYEP BIZEW QYPKZ FHAAM GWPTR XNYWR LKSQE XHGRA

QCWAV JNCPM HDHZT BCBHR AMXUE OLTWR RIKNQ AKKDZ VJOYW
                                                    bet
WHQJR FGYVP GILWV WGPTF MLYKX TAKOZ ATFGL AUT
WEENL ATESE PTEMB ERAND DECEM BEROF THATY EAR
```

We now have a slightly different way of finding the cycle if we need to employ it, but first, are there any significant repeats?

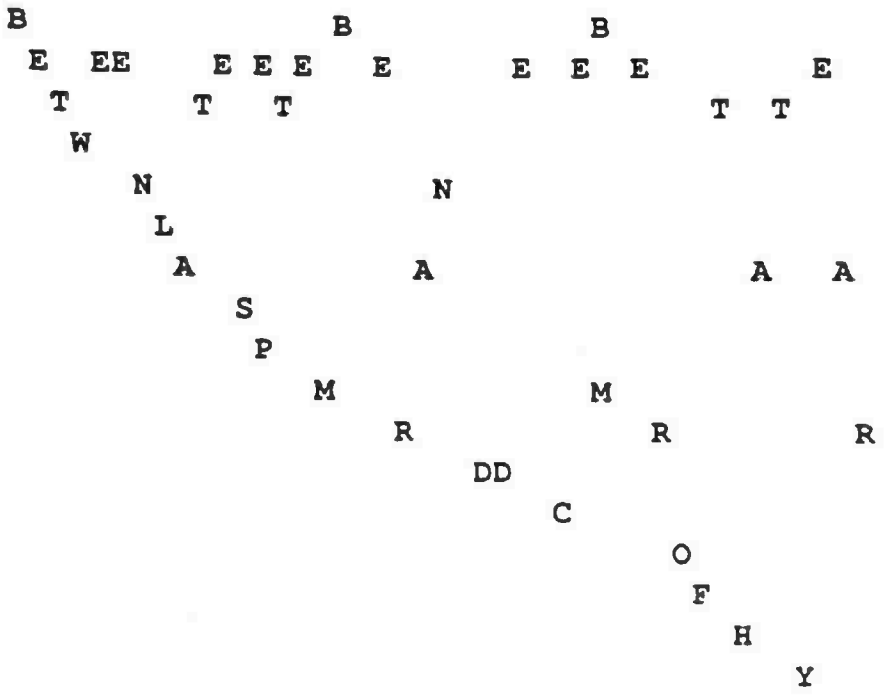
The WQY may not prove valid, with HEG a so-so plain, since it can not be THE. (The plain T would be already spoken for, as cipher M.) In any case the interval is 58, which is highly discouraging, with no sensible factors.

Let's look at the cribs so generously donated. Obviously a plain hit at the correct interval of the cycle would result in a cipher coincidence at that same interval. Similarly, with two occurrences of a plain letter at some interval other than the exact cycle (or multiples of it, of course) the ciphers cannot be the same.

(An exception is possible in the event of two rows of the tableau being identical, because of the repetition of a letter in the "vertical" key — the word used to determine the various settings of the cipher alphabet.) Looking more closely at the pairings at the end:

```
OYWWHQJRFGYVPGILWVWGPTFMLYKXTAKoZATFGLAUT
betWEENLATESEPTEMBERANDDECEMBEROFTHATYEAR
```

A tricky way of looking for repeats, if you've got plenty of paper, is to write the plain on different levels for each letter value:



Indexing these repeats according to the distance (1 being "next door" and 2 being removed by one letter, etc.):

```
1 E, D
2 E
3 T
4 E, A  etc.
```

You must remember to count longer intervals which have an intervening hit, such as E--EE---E-E, which in addition to the obvious 1, 2, and 4 has a 3 due to "bEtweE" and an 8 due to "bEtweenlatE, and so forth.

But the simplest eye inspection of the cribbed-in areas may show repeats of the plain-cipher pairs, which give an immediate clue as to cycle length. In case there are no repeats you can use negative information: consider repeats of plain which do NOT result in a hit, and knock these intervals out.

Thus, Y over E, and H and Q over E several positions away "knock out" the intervals 3 and 4 (which you wouldn't want anyway) and also 11 because of the Y over P. But note the hit we have been looking for — Y over E, and Y over E 9 places later. Written out on this cycle:

```
MWQYDKMCA OKHSEEYUL IHWYTEWYR LHGLMEJCZ HAKENYWUP
THEGREATE             E GH EE   E E           E A

QSQSOESYE PBIZEWQYP KZFHAAMGW PTRXNYWRL KSQEXHGRA
  E   hE  E T    EA R      RT ER    E   R E     E
```

```
QCWAVJNCP MHDHZTBCB HRAMXUEOL TWRRIKNQA KKDZVJOYW
   T    A TE                  NH  E   E R     BET

WHQJRFGYV PGILWVWGP TFMLYKXTA KOZATFGLA UT
WEENLATES EPTEMBERA NDDECEMBE ROFTHATYE AR
```

You are naturally pleased with the confirmations, and may safely write out the tableau:

P:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
C: 1	U				P									T				K		M			W			
2				F	H			W								O	G	T								
3				M	Q	Z														I						
4					L		Y							J						A						
5			Y					T				R	W					D								
6	F	V			K																					
7	M	O			W								X							G						
8		T			Y													G		C					L	
9	P				A		C	(See	below)										V	W						

Now comes the hard part. Since the alphabets are different, we cannot chain from the plain to the cipher. However, WITHIN the cipher, the same rules apply as before — except there isn't nearly as much information. Let us look for pairs on the Cipher 1 row and see if we can pick up something on other rows: for instance, U to P is the same distance as F to K, M to W and P to A. OK? Let's see what we can dig up. Remember that we are dealing with unknown decimations, and the relationship between the UPA, the PT and the PK below is unknown.

```
UPA FK MW      Plain A to E and Rows 1 to 9
PT LJ             "  E to N
PK HT YG          "  E to R and Rows 1 to 6  (adds UF)
PM QI LAWG YC     "  E to T and Rows 9 to 7 and Rows 4 to 9
UMG PW            "  A to T and Rows 1 to 7
TM JA             "  N to T
FH MQ             "  D to E
WTD               "  H to R and Rows 2 to 5
FV MO             "  A to B
VK OW TY          "  B to E
OG TC             "  B to T
PH KT          Rows  1 to 2
PQ MI                1 to 3
PL TJ MA             1 to 4
PY KG MC             1 to 8
FM HQ KW VO          2 to 0  and 6 to 7
HY TG                2 to 8
QL IA                3 to 4
QW IG                3 to 7
QY IC                3 to 8
QA IW                3 to 9
LW AG                4 to 7
LY AC                4 to 8  and Plain A to G (adding
                             cipher C under Plain G on Row 9)
FP KA                6 to 9
OT WY GC             7 to 8
YA CW                8 to 9
```

You will notice I combined a few relationships, such as the Row 2 to 3 and 6 to 7 additions. Going over this tableau with a finetooth comb might reveal other additions I didn't bother with, but you get the idea, and we will proceed together.

Other hints may be given, such as the fact that S and T in the plain sequence are likely adjacent (from the VW of Cipher 9) but these are slim pickings.

Some of the partial sequences look good, such as FH and MQ, with or without an intervening letter.

But it's the LAWG which interests us, for two reasons. It has the longest partial chain, and it ties in plain E and T at the same decimation; we know this decimation is not the final one, but let's make a partial tableau using what we know:

```
Plain:      E T

Cipher:     P M
            H
            Q I
            L A W G


              K
          L A W G
              Y C
            L A W G
```

Notice that if our FH and MQ order is right, P must be in the keyword, since the reverse bits of the above (MP, IQ, GWAL) would not be consistent with MPQ.

Since we have run out of gas, it would be good if we could guess a little more plain. How about that bit of plain that goes E-GH-EE? There seems to be only one good answer, even assuming words that end with E-G or begin H-E.

Why not try EIGHTEEN? And since they are talking about "years", why not put in SEVENty, since the Es are so fitting? And the plain T of SEVENTY is confirmed.

You understand that in some problems with more cipher to work with, we would be able to push out more plain text on the basis of a modest "guessing" effort. But in this case, we can use the information we get, not so much to read more plain as to assist in the recovery of the sequences. The plain V of SEVENTY may not produce much more text, but the placing of the cipher G may be a bonanza.

Continuing with the LAWG section we were working on, these new values add KE and JR to the chain, so we put them in the tableau.

```
MWQYDKMCA OKHSEEYUL IHWYTEWYR LHGLMEJCZ HAKENYWUP
THEGREATE      T     EIGHTEEN SEVENTY          E A


QSQSOESYE PBIZEWQYP KZFHAAMGW PTRXNYWRL KSQEXHGRA
  E   ThE   E T    EA R      RT ER    E   R E     E


QCWAVJNCP MHDHZTBCB HRAMXUEOL TWRRIKNQA KKDZVJOYW
   T    A TE                  NH  E   E R     BET


WHQJRFGYV PGILWVWGP TFMLYKXTA KOZATFGLA UT
WEENLATES EPTEMBERA NDDECEMBE ROFTHATYE AR
```

Plain:	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Cipher 1	U				P									T				K	L	M			W			
2				F	H			W							?	G		T								
3				M	Q	Z			W											I		G				
4					L		Y							J						A						
5			Y					T				R	W	M				D								
6	F	V			K													J		E						
7	M	O			W								X							G						
8		T			Y													G		C					L	
9	P				A									R					V	W						

Time for a look at areas which look good for the final version of the cipher alphabet, such as the VW and LM we pointed out earlier on.

The latest tableau has KLM under the plain of RST. (Remember, this does not say these trigraphs are in the SAME alphabet, but only that if the cipher sequence has KLM, the plain has RST, both suitable candidates for consideration.

If we can push the "final" version, where it is easy to keep track of partial sequences that "look good," we can abandon the decimation LAWG, etc., which moved us along a bit, but may have served its purpose.

Note that if KLM is correct, then G-C is correct. Rows 7 and 8 have a G and C under plain T, and WY under E (and OT under B). Looks like we have hit pay dirt, for this means that WXY and O-T are part of the final chain. Let's go back to the drawing board and push these partial chains that we like: KLM, G-C, VWXY, EA, O-T, and other good-looking bits as we find them.

Fiddling around with possibilities, it looks as though the cipher sequence goes JKLMQVWXYZ. It sounds like a leap into the wild blue, but if we are close it will probably pan out:

A	N	D	E		I	C	B	F	G	H
U	T		P							
		F	H	J	K	L	M	Q	V	W	X	Y	Z
J	K	L	M	Q	V	W	X	Y	Z				
		J	K	L	M	Q	V	W	X	Y	Z		
J	K	L	M	Q	V	W	X	Y	Z			T	
		F	H	J	K	L	M	Q	V	W	X	Y	Z
J	K	L	M	Q	V	W	X	Y	Z				
J	K	L	M	Q	V	W	X	Y	Z				
P	R		A							

It's amazing how the thing snowballs, and if you make a correct assumption the total answer is quickly there. Of course, you have to be prepared to grab a fresh sheet of paper and start over if you stray too far afield, but if you keep scribbled (but accurate) notes of where your branch points of reasoning are located, it is well worth a fishing expedition.

You can finish recovering the text and sequences. By now you will appreciate the fact that without a clue as to the plain text, a problem such as this can be very, very difficult. Fortunately you are likely to encounter problems that the "cryptographer" intended you (as the "cryptanalyst") to solve without too much difficulty.

In a "real life" situation such a simple system would probably not be used, but if it were, you can be sure there would be no cooperative attempt on the part of the originator to make things easy for you.

Incidentally, don't forget to look for the keyword that controls the cipher setting. Oftentimes you can get a lot of information when you note that under a certain plain letter you have cipher letters that read -E-UB-I-AN or some such recoverable combination.

In coping with QUAGMIRE IV, one often can make a relative solution and recover the plain text without having the true order of the sequences show up as obviously as in the above example. If you have most of the 26 letters in your relative sequence, it is usually an easy job to try the various decimations and find the correct order.

In any case, the procedure is similar to that demonstrated in monoalphabetic substitution with K4. There is still just one plain alphabet and one cipher alphabet to recover.

Chapter VI

DIGRAPHIC SUBSTITUTION

IT WOULD be too much of a good thing to describe all the various low- and-medium level cryptograms one might encounter in the ACA publication, puzzle magazines and spy novels, etc., but several of them will be selected as being representative.

One of the categories, probably next in difficulty to the monoalphabetic versions previously described, is Digraphic Substitution. Of course, like many of the other problems discussed, they can be difficult if you have no entry point, such as a specific crib, or if you don't have at least an idea about the kind of text you have.

As the name implies, Digraphic Substitution replaces each plain digraph with a cipher digraph, and in many systems consistently so, in that if TH is replaced by OG in one part of the message, you can be sure the OG cipher is plain TH wherever it occurs. (Warning — some systems don't have this consistency.)

Playfair

And of the Digraphics, possibly the most historic is the Playfair, which although bearing the name of Baron Playfair of St. Andrews was actually invented by Charles Wheatstone. This latter Victorian gentleman is better known for his electrical resistance measuring device, the Wheatstone bridge, and (at least to cryptologists) as the inventor of a cipher device.

The Playfair system utilizes our old friend, the 5 x 5 square, in which a keyword-mixed alphabet is transcribed (omitting J in English, or perhaps W in French usage, for example).

```
N O W I S
T H E M C
A D F P R
Y B G K L
Q U V X Z
```

Each plain digraph is replaced by a cipher digraph from the square according to the rules mentioned in reference to XIMENES puzzles (in Part I): If you visualize an imaginary rectangle of which the first letter of the digraph is the one corner and the second letter the diagonally opposite corner, you take the remaining corners as your cipher, moving up or down (not left or right) from the position of the plain to get your first cipher letter. (Easier done than said.)

Thus TL becomes CY (and CY becomes TL — this reciprocity being a very important factor in solution.) Similarly DI is PO. If the two letters of plain are on the same row,

one reads the letter to the right of each as a member of the cipher digraph. If letters are in the same column, take the cipher underneath the plain value.

If you suddenly run into a double letter to encipher — SSIO or TTIN or something — don't panic. Just throw in an X, and encipher SX SI or TX TI and continue on your merry way.

Thus CH is TE — the square being considered cyclic in all directions. HE is EM (and this "shadow" of the second plain being the first cipher in such cases is an important feature in solution.)

Although the CRYPTOGRAM may be kind enough to label the system, and present the cipher in digraph form, in "real life" the message would probably be sent in groups of five letters. How would you diagnose it?

Suppose we had a message:

```
VRDFU XTFWU EAOCH OUHLU HECBC ZREGP IBUSP LPOPL PBUSD ROFGP
KUUZN ADTZB HOGPL HLAAB ZLFZL ZUMKD UZFKC TOIBZ COTKQ GALHO
LEHOR BSCEM US
```

What can we gather from the cipher? First, there are no J's, implying a 25-letter alphabet and probably a 5 x 5 square. (If half the I's were changed to J, it would probably still stand out, since the sum of I and J would be roughly equal to the average cipher letter's count.)

The monoalphabetic frequency (apart from J) may be pretty flat, if the square is carefully constructed, in terms of measurement that remain to be discussed in further chapters. The digraphic count should be rough.

Now an important count: How many repeated letters do you see as a digraph — AA, BB, etc.? I see two, unless I missed one in my cursory inspection.

How many would you expect in a message of 112 letters? Since one out of every 25 letters is a probability for a repeat with a 25 letter alphabet, we should expect between 4 and 5 — not a highly significant discrepancy, but still interesting.

More specifically, where do these AA pairs appear. One begins at the 52nd position and the other at position 68 — in other words "off the cut" if we counted 1-2, 3-4, 5-6, etc. No digraph, marking them off from the beginning, would have an AA pattern.

This is a slightly more impressive fact than the 2 occurrences within the entire text, because in digraph form, with 56 pairs to count, we should have more than two of the AA pattern and we get none. This isn't "statistical" reasoning, but one is heartened by the fact that, although it doesn't HAVE to be Playfair, if it WERE Playfair it would have this property.

So maybe Playfair? In the actual case I'm giving you, the problem gave a crib "THEOTHERPERSON". (It also had a title — "Politeness".) Now the cut of the digraphs of the crib could be either way, TH EO TH ER PE RS ON or T HE OT HE RP ER SO N.

In either case, there is a repeated digraph with one pair intervening.

Looking at our message written out in digraph form, we note PL PO PL and HO LE HO, so let's examine the possibilities of general placement with two possibilities each for the hitting digraph:

```
US PL PO PL PB US DR OF        AL HO LE HO RB SC EM US
   TH EO TH ER PE RS ON           TH EO TH ER PE RS ON
 T HE OT HE RP ER SO N          T HE OT HE RP ER SO NX—
```

Can a plain letter have its cipher equivalent in exactly the same position — E over E, for example? Not in Playfair. So the first possibility of the four above is knocked out. By the same token, the HO over HE is impossible, so we have two places to try.

Moreover, the US going to -T and also to ER in the second possibility of the first placement is patently impossible, so we only have one possibility for exact placement.

```
VR DF UX TF WU EA OC HO UH LU HE CB CZ RE GP IB US PL PO

PL PB US DR OF GP KU UZ NA DT ZB HO GP LH LA AB ZL FZ LZ

UM KD UZ FK CT OI BZ CO TK QG AL HO LE HO RB SC EM US
                                  TH EO TH ER PE RS ON
```

Remember what we said about the "shadow" of some pairs which appear in the same row or column? (By chance, eight times out of 24 the second letter of a plain digraph will be in either the same row or the same column as the first letter. When this happens, there is a 50-50 chance that either the first letter of the plain digraph will appear as the second letter of the cipher digraph, or the other way around.) Given PO over OT, we know that either a row goes TOP or a column has those letters going down.

Also, the PB RP pairing shows that RPB must have the same property. Since O leads to P in one instance, and R goes to P in the other, they cannot all fit on a row or column, but must represent one of each:

Now we can add a few more letters in the "relative" square we are trying to build up, with HO = TH, LE = EO, RB = ER, SC = PE, EM = RS and US = ON.

Remember you may be going off course if you ignore the possibility that the two plain are in the same row or column; but it usually pays to assume at first that all letters used are at corners of the rectangle made by a down-up-down sweep, or an up-down-up movement from the first letter of the pair.

If you are lucky enough to have information that proves letters to be in the same row or column, so much the better.

Also, be careful not to put letters into a "known" column without proof, adding pairs as you get them without regard for the 5 x 5 limitation until you get some sort of check.

We know that THO goes together on a row or else in a column; also OEL has this property, and ERB. (It is rare that three such goodies all fall within the space of one crib.)

The OEL and ERB permit of only two possibilities, and we can add in the THO in two possible ways, giving four tentative arrangements to which we can add other letters on the basis of information contained in the crib:

```
  T H O              T H O E L
c   E R B              R C
      L                B
P   s m                  P
                         S

      T                  T
      H                  H
      O                c O E L
      E R B              R
      L                  B
                       p   s
```

But we get into trouble on three of the possibilities: We can add part of the chain into the second "square", but by the time we come to RS = EM, we find that there is already a letter C where we need the E (or saying it a different way, the E and the R are on different rows when they can't be.)

In the third square, we can't even add PE = SC, since it would result in too many letters on a row or column any way we try it. And the fourth possibility bogs down for the same reason that the second one did.

Hence our square has to be (relatively):

```
    T H O     u
c       E R B
        L
p       s m   n
```

Notice the the cipher HO UH LU HE deciphers to TH ?T ?O O? I'm not saying that you would immediately try THAT, but sooner or later you would, and you could test the assumption:

```
    A U T H O
c R B       E
            L
p m n       s
```

You undoubtedly will see that you almost have placed (by accident) the square in its primary position. (The small letters can be shifted if necessary, but the capital letters are fixed.) This is one of the nice things that occasionally happen, in the sense that you could have made the square

```
n     s m p
U T H O A
B     E R C
      L
```

and not become particularly interested in keyword possibilities. In any event, if you try the keyword AUTHOR, you will have no trouble deciphering the entire message with the square:

```
A U T H O
R B C D E
F G I K L
M N P Q S
V W X Y Z
```

Again, you may remark on the fact that in some instances things just fall into place, and the rest of a solution is merely clerical work. You will soon realize, however, that the other side of the coin dictates that you will tackle a problem which seems just as simple as the one given, and spend much more time beating it to death.

Because Playfair is so important a system, and much more "fun" than some ciphers, let's try another. The title is "Electronic woes" and the crib OMPUTERROOMNOON.

```
CT NY KH UA GP YI SX PG XY NM PC TB RY NW EM PF SQ EM PM
KY CL NM YP QS BP EZ SF QA LZ RY SM GU FE TB RY NW EM QF
QY TU GB CP CT GU EY YI YQ YI MQ GP YI TD RY NW EM QZ NP
RY OP PO TD GU KY FN BV SA FP SD EZ
```

The RR in the crib would definitely have to be split, either on the cut UT ER RO OM NO ON or by having an X thrown in, as would have been necessary with PU TE RX RO OM etc. In either case, part of the text runs RO OM NO ON, with the pair NO ON having to fit under an AB BA pattern. Fortunately there is only one such — OP PO. (These reverse pairs are almost as important as repeats.)

When you look at the cipher in conjunction with the two possibilities of text (with and without the X between the R's) you find that there is a repeat of RY at the spot where the OM of COMPUTER fits, so there is no ambiguity.

```
YI TD RY NW EM QZ NP RY OP PO TD GU KY FN BV
      CO MP UT ER RO OM NO ON
    C OM PU TE RX RO OM NO ON
```

Also NP deciphers as RO, which means that in all probability the proper sequences goes NOPQR. A word of caution. Some squares are created by spirals. There is even boustrophedon movement, or reading a row forwards and the next row backwards, then forwards, etc. (This is an unusual word that comes from the Greek meaning of "turning the ox in plowing.") Therefore the NOPQR might be a column.

The plain OM cipher RY pairing will not work on a NOPQR column sequence however. (Try it.) It must therefore be a row.

When you add the pairing of NW=PU, there is a good chance that you are half-way home, since the last row of the square is probably UVWXZ. And TEM could either go on a row or a column, yielding two possible squares:

```
          T
          E
    Y     M          Y T E M (Row 1, 2 or 3)
N O P Q R          N O P Q R
U V W X z          U V W X Z

CT NY KH UA GP YI SX PG XY NM PC TB RY NW EM PF SQ EM PM O
   O                    V? R        OM PU TE
   O                    VE R        OM PU TE
```

Hold the phone. When you start to decipher, and come to COMPUTE, isn't it worth a chance that the word is COMPUTER rather than COMPUTING or COMPUTATION or whatever? Especially since P is on the same row as R. So you can eliminate one square possibility, and put in the F. But here I quit, since a little thought will certainly enable you to fill in the entire text.

The only thing wrong with what we have done so far, is to possibly mislead you into thinking that all Playfairs are this easy. I'll put one in the section at the end which will possibly try your skill a trifle more exactly.

Fractionation

An interesting type of cipher results from a process known as fractionation. Each letter (usually of a 25 letter alphabet) is given a two-position designation (usually two numbers within the 1 to 5 range, representing coordinates on a square.[1]

```
    1 3 2 5 4
  ┌──────────┐
1 │ F R A C T │
3 │ I O N B D │
2 │ E G H K L │
5 │ M P Q S U │
4 │ V W X Y Z │
  └──────────┘
```

Often the correspondents (by prearrangement) use the normal 12345 for coordinates, but they may be mixed, as in the example. One can visualize the not-too-sophisticated cryptography of using the same square several times, but with different mixed coordinates.

(You will note that the row and column designations are the same sequence, 13254 in this case, though this may not be absolutely adhered to.)

"A" is represented here by 12. Each plain letter is replaced by the proper dinome, and the string of resultant numbers is then manipulated in some fashion, and these numbers (presumably in a different order within the "string") are then recombined (using the same square as in encipherment, so that 12 represents A.) Use of a different square for recombination would provide a lot of extra security, as one might guess.

In a fractionation system, the above "manipulation" involves the separation of each plain dinome into its two halves, and recombining them after some manipulation, such as encipherment, transposition, or whatever.

Get the plain value as a dinome,[2] separate the two halves in some way, manipulate by encipherment, reading in a different direction, or whatever way is stipulated, and then recombine to get a cipher value. (Naturally the method has to be capable of easy reversal, to get back from cipher to plain.)

We will examine some of the ways this manipulation can be performed.

Bifid

The CRYPTOGRAM has developed a few variations on this theme, the most common one being called a Bifid. The plain text is written out in some width (usually between 5 and 10) and values from a square (such as the one above) are placed underneath the letters as a vertical digraph, so that the first row of numbers represents the row names and the second row represents the column names.

Thus the text "This is an example of", written out on a 7-wide format, would appear:

```
T H I S I S A     N E X A M P L  E O F ....
1 2 3 5 3 5 1     3 2 4 1 5 5 2  2 3 1
4 2 1 5 1 5 2     2 1 2 2 1 3 4  1 3 1
```

The figures in the first 7-wide set are now read from left to right, beginning on the first row two at a time, as coordinates on the same square, to yield cipher. You will note that the fourth cipher will be 14, from the last of the first line to the first of the second.

This gives 7 letters — the same as the number of plain letters on this width. The identical treatment is given the second series, the third, and so on, and the entire cipher text is then transmitted in 5-letter groups. The cipher for the above plain would be:

```
A B B T E   M Q O V S   H A E D G  etc.
```

Shall we try to work through a Bifid? The title is "Speed record" and they tell you that the message ends DOWN IN THE OCEAN. In this case we are given the width as 9, but they are not always that generous — and in another problem we'll have to find the width out for ourselves. Also, since we are trying to get you familiar with

techniques, as much as having you solve plain text, a little extra help in the crib would not be amiss. So we will tell you the message starts IN NINETEEN.

The cipher as we see it is usually written out in five-letter groups, K D R U O Q M T Q N H N P L D P Z V A B G G Z K I, etc. But only occasionally is five the true width, of course, and we know it to be nine in this case.

Writing the cipher text out 9 wide involves placing 5 letters on the top row and 4 letters (offset) on the bottom row, if you visualize in your mind's eye how we went from the string 1 2 3 5 3 5 1 4 2 1 5 1 5 2 in the above example to the cipher A B B T E M Q:

```
I N N I N E T E E   N
K   D   R   U   O | N   H   N   P   L | A   B   G   G   Z
                  |                   |
  Q   M   T   Q   |   D   P   Z   V   |   K   I   F   A

L   A   A   S   B | B   G   E   Z   T | S   R   D   D   F
                  |                   |
  G   L   C   T   |   L   O   E   K   |   V   P   B   C

C   Z   C   G   N | K   G   A   W   I | G   O   H   S   D
                  |                   |
  B   E   O   A   |   B   B   P   E   |   Q   I   Q   B

D   B   F   R   X | G   E   A   H   I | E   S   B   T   O
                  |                   |
    Y   S   M   R |   T   H   Q   C   |   D   K   Y   X

                                  D O   W N I N T H E O C
B   W   S   E   S | X   G   C   Z   A | Z   K   B   C   O
                  |                   |
  B   L   F   W   |   I   C   B   I   |   V   T   L   P

E A N
C   N|
     |
  L  |
```

Note that the last box can be any number of letters, from one to seven. I have placed the proper plain text of the crib over the correct position of the cipher.

Now is a splendid opportunity to illustrate a weakness of this system, which we exploit at every opportunity — placing a crib, guessing plain text, etc. You will note the last cipher is N, and the last plain is also N. Looking a little further, in the entire crib, beginning and end, we have 14 cipher letters on the top row underneath plain text of the crib. What are the chances for a hit with 14 comparisons, and a 25 letter alphabet? Well, we might expect none, or one, or possibly two hits by random. Indeed, we get two. But look as well at the cases in which a letter on the bottom row matches the plain

letter one position to the right. We get two more hits, with only 10 opportunities to count them (without getting involved in a specialized case of "hitting").

This makes a total of 22 comparisons, with four hits. Remember that "accidental" matches between random streams of letters (which these are not) will happen only once in 25 letters, on the average.

If we had the entire text of 138 letters put above the cipher, 72 of them would be over cipher letters on the top row. By chance we should get about 3 hits.

Would you be surprised to learn that in this system we would be more likely than not to get about 14 — a few more or less?

The reason for this "hitting" is as follows: Consider any position. (Take the W of DOWN over the cipher Z, for example.) Is it clear that whatever ROW name W has, Z has the same number? Similarly I and K have the same row name, and so on.

Remember that the CIPHER, in terms of the two-digit number, is read from right to left (and progresses from the last position on the top row to the first position on the bottom row.) But the PLAIN is written vertically, from any given position in the top row to the position directly under it.

Now why didn't we get a cipher W instead of Z, or cipher I instead of K? Because the W, going down, might be 12. while the cipher that starts in the same place, but reads to the right, depends on an entirely different plain letter for its value.

In this case, the plain N of DOWN is directly above the position we are discussing.

If N begins on Row 3, our cipher will be 13; if N begins on Row 5, our cipher will be 15. and so on. But there are only 5 numbers in our system — 1 to 5. Therefore, the chance is one out of five that the second position of the cipher stream at the point under discussion will be a 2.

If it is, the cipher under the W would read to the right as 12, the same as the plain W going down. And we would have to get a W cipher at that position.

Now consider the values on the bottom row of our write-out. The same general reasoning prevails. Look at the V under the N of DOWN. Isn't it true that the COLUMN number of V is the same as the COLUMN number of Plain I of DOWNIN?

Therefore, if Plain W were 53, for example, and I were 3-anything, the cipher on the bottom row under W would have to be I.

It does not matter whether I is 31, 32, 33, etc. — the only important point is the answer to the question "Is the Column name for W the same as the Row name for I?"

Summarizing: Looking at any given letter on the top row of the Cipher stream, the chances are one in five that the Plain above it will match. And looking at any Cipher letter on the bottom row, the chances are the same one in five that the Plain value IMMEDIATELY TO THE RIGHT will match.

Not bad odds, 1 in 5 as opposed to 1 in 25 by "chance". If this isn't perfectly clear, play around with pencil and paper, enciphering, deciphering — and you will see how often you say "Oh, I see. If this is Row 2 Column 4, the other has to be 24 as well."

The above phenomenon occurs when the width is odd — 7, 9 etc. If the width is even (6, 8, etc.) a similar phenomenon occurs, but in addition, a digraph in plain frequently shows up as the same digraph in cipher. (More of this anon.)

Back to solving our message. We start with a blank sheet of cross-section paper, and no preconceived notions of the inside of the square or the numbers of the rows and columns. Our only rules are:

(1) The numbers of the rows will follow the same order as the numbers assigned to the columns.

(2) We can assign any value we want to the the first (or second) position of the first letter we decided to tackle — whether it be a plain or a cipher value.

But having assigned that number — for example "1," we must be careful not to make an arbitrary assignment of that number any place else. We can call our next assignment 2, or 3, or 4 , or 5, but having used up that "opportunity" we must be careful not to use the same number unless it is dictated by the first or second assumption.

(3) Any square we recover from this method, or the names of the rows and columns, will usually be on a relative basis. It will work, and decipher letters correctly, but only in the later stages will we determine the exact positions — though we should utilize any information pointing in this direction.

So let's call some letter something. The best bet is to find the letter occurring most often (counting both plain and cipher together) and to assign it a value of 12.

```
I N N I N E T E E   N
K   D   R   U   O | N   H   N   P   L | A   B   G   G   Z
  1 1   1         | 1 2     1 2       |
  Q   M   T   Q   |   D   P   Z   V   |   K   I   F   A
  2 2   2         |   2               |

                                      D O   W N I N T H E O C
B   W   S   E   S | X   G   C   Z   A | Z   K   B   C   O
                  |                   |   1   1
  B   L   F   W   |   I   C   B   I   |   V   T   L   P
                                        2   2

E A N
C   N
    1
  L
2   2
```

Look at the beginning and ending spots, where we have placed the crib. In this case we have called N 12, since there are so many plain N's, and a few cipher N's as well. Note that plain N's read 12 going down, and cipher N's read 12 going across (or from the last top to the first bottom when they fall in the last position on top.)

Notice that D and R must also be on Row 1 — an arbitrary designation, of course. But we would be safe in calling D 13 and R 14. (Note that we cannot call either of them 11 until we get evidence to that effect — 2 MIGHT be 1, or 3 MIGHT be the same as 1, but we cannot take any chances, yet.)

On the other hand, filling in N as 12 we find Q to be 22. If we used 2 as our point of departure rather than 1, we would be perfectly safe in calling certain things 3, 4 or 5, as long as we don't call them 2 without some justification.

For instance L is -2, so let's call it 32. Let's start our relative square, and fill in what values we get:

	1	2	3
1		N	D R
2		Q	
3		L	
	K		
	Z		

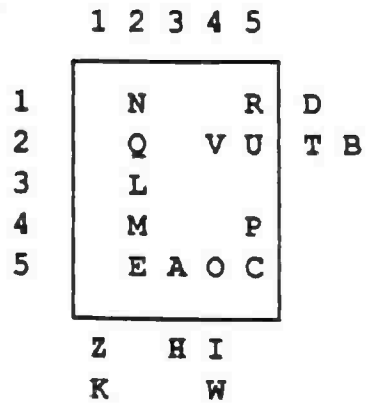
Note that K and Z are both "something" 1. Let's work just with the first and last areas, and push each value forward as we get it.

```
I N N I N E T E E   N       R     E
K   D   R   U   O | N   H   N   P   L | A   B   G   G   Z
  1 1   1 5 2 5 5 | 1 2   3 1 2 4 5 3 |   3 2
  Q   M   T   Q   |   D   P   Z   V   |   K   I   F   A
4 2 2 4 2 2   2 2   2 1   4 5   1 2 4   1   1   4     5 3

                            C E   D O   W N I N T H E O C
B   W   S   E   S | X   G   C   Z   A | Z   K   B   C   O
2     4     5 2   |         5 5   1 5 |   1   1 2   5 5 5
  B   L   F   W   |   I   C   B   I   |   V   T   L   P
  2   3 2       4   3   4 5 5 2     4   4 2 4 2   3 2 4 5

E A N
C   N
5 5 1
  L
2 3 2
```

Here is the square as we have it so far:



You realize why we called E 52 — when we discovered it ended in 2 (because of the 22 value for Q) there was no other place in the square for it.

There are so many things we can choose to do, as the solution snowballs, that we have an embarrassment of riches. We can look for places where the plain text is becoming more obvious, or we can play around with the square, hoping to put it on the correct base — in which case we'll just go ahead and read the message.

We can also look for repeats — for you can see that a repeat of a digraph formed by a letter on the top row and its lower neighbor next door must give a repeat, either fore or aft. The solver will quickly learn how to take advantage of all these hints, and will independently discover many more ways of improving the chance of success.

Small hunches often lead to bigger developments, so I wonder what would happen if we made the logical assumption that there was a past tense verb in front of "Down in the ocean." Do two small hunches add up to a good guess?

For in many squares (even those with unusual inscription methods) Z will be the last letter, and either 55 or 33 in a 1-2-3-4-5 designation. Note when we assume CED, we get the bonus of an XX pattern for Z, 11 in the base we are using.

How about FORCED DOWN IN THE OCEAN? Go ahead and try it. (The 4 and 5 in the bottom row partially confirm this.)

Also, we can straighten out B and D if this is right. Notice that B can be only 21 or 23, while D is 13 or 14 (Z being called 11.) Obviously the answer is that B is 23 and D 13.

```
I N N I N E T E E   N     I R     E I   G
K   D   R   U   O | N   H   N   P   L | A   B   G   G   Z
3 1 1 3 1 5 2 5 5 | 1 2   3 1 2 4 5 3 | 5 3 2 3 5 1 5 1 1
  Q   M   T   Q   |   D   P   Z   V   |   K   I   F   A
4 2 2 4 2 2   2 2   2 1 3 4 5 1 1 2 4   1 3 1 3 4       5 3
```

(Showing beginning and end only)

```
                                F O R C E D D O   W N I N T H E O C
B   W   S   E   S | X   G   C   Z   A | Z   K   B   C   O
2     4     5 2   |     5 1 5 5 1 1 5 | 1 1   1 2 3 5 5 5
  B   L   F   W   |   I   C   B   I   |   V   T   L   P
  2   3 2       4   3 3 4 5 5 2 3 3 4   4 2 4 2   3 2 4 5

E A N
C   N
5 5 1
  L
2 3 2
```

The good fortune is apparent, so you can finish this all up, including recovery of the square. (Those of you who have been champing at the bit, and saying "how about that date?" can now go to work.)

Twin Bifid

A slight twist on the type of problem you have just solved consists of a "real-life" situation in which two messages start out with the same plain-text for a stretch, and then each one takes off on its own.

Since common beginnings are a well-known entry point for many otherwise difficult situations, the "Twin Bifid" (as the CRYPTOGRAM labels it) serves as a lesson of "what not to do", but is also good fun to solve.

Actually, it may be easier than a single message which contains a crib. (They give you no tip with a twin, since it almost solves itself.)

Usually they will say "First 23 the same", or something like that when such a problem occurs, but instead I will pretend that Zendia[3] has a stupid cipher clerk who sent the same message to the Ambassador in Minnewawa that he sent to the Consul in Oohaha, but used a different width for the encipherment.

So the entire message says the same thing. (Logically, it would have a different message number and perhaps a different "salutation", but we won't bother with such trifles.)

Here are two texts:

```
Msg. No. 1:  FQCIU KIPFF EOZZS FQTRV DOWUH RBEMN HZIK
Msg. No. 2:  FQCZY NBXEP DEOFZ SRXFW YARDL UEMBH ZNIK
```

One can immediately see that the messages have a lot in common. A particularly significant fact is that (considering the messages placed one under the other) repeats are found both in the same position and slightly removed. There is also a feeling of "in and out" or "hit and miss."

Diagnosis of the phenomena displayed is probably dependent on one's having experience with certain kinds of "busts", resends, near-dupes, and other types of messages the cryptanalyst is continually looking for.

The ending in particular is interesting, with the bottom message having an "N" extruding in the middle of what would be a four-letter repeat, HZIK — and strangely enough, this N appears just ahead of HZIK in the first message.

How does one write the two messages out so that the repeats at the end would become more consistent — and perhaps cause repeats in other areas to make a trifle more sense?

One familiar with the Bifid might sense that an element of 3 entered into the cycle of one or the other of the messages, in order for the repeat at the beginning to break off at that point. Of course, this implies a 6-wide cycle on a Bifid base — and the text of both messages might be written out and compared against the other message, with a couple of groups written at intervals of 7, 8 and 9.

At some point in this seemingly trial and error method, the first message (6-wide) would be compared with the second (7-wide):

```
#1 | F  Q  C  I  P  F  Z  Z  S  R  V  D  |
   | ------- |        |  ------|--       |
   | I  U  K  F  E  O  F  Q  T  O  W  U  |
                 -------                 --
#2 | F  Q  C  Z X  E  P D Z  S  R  X  A R
   | --------|           |--------    |
      Y  N  B |  E  O  F|  F  W  Y   |  U
                 ------                --

#1  H  R  B  H  Z
          -- |------
    E  M  N  I  K
    ----------------
#2    D  L B  H  Z
          |----------
    E  M  N  I  K
    ----------------
```

Now it becomes apparent that the plain must be exactly the same, and the difference arises when the cycles of 6 and 7, respectively, cause one of the two streams to be listing a column number while the other stream lists the same number as a row.

The numbers are the same, but the letters at times may be different, since one may be reading Row-Column while the other is reading Column-Row at the same point.

You will see that the endings have to be written at whatever width is necessary — four wide for the first message, six wide for the second.

How to attack? Obviously we get no further information when we assign a precious number to a repeated letter such as cipher C for the 5th plain letter. (It isn't even repeated later on in the message.)

The ones that "pay off" are those which are offset, such as the I in the first message beginning. And this starts a chain, which we shall pursue.

Assign a value of 12 to I, and you will see that this automatically makes Z end in 1, and Y begin in 2 (if we are going to make the messages read with identical numbers in the streams.)

Keep on assigning new numbers as required, building up the square, and putting in all cipher and plain values as they occur, making sure to match the two versions with any numbers you enter.

You will quickly find that Z is 11, which makes it safe to add 13, 14 and 15 when and if you can progress in that direction. You will find that E is the next letter to fit on the first row, so assign it 14.

```
                        E   P E     E  etc.
#1 | F   Q   C | I   P   F | Z   Z   S | R   V   D
   | 4 3 3     | 1 2 3 1 4 3|1 1 1 1 etc|
   |           |            |           |
   | I   U   K | F   E   O | F   Q   T | O   W   U
   | 1   2     | 4 3 1 4    |4 3  etc.  |
```

```
#2 | P   Q   C   Z|X   E   P   D|Z   S   R   X|A   R
   | 4 3 3       1|2   etc.     |             |
   |   Y   N   B  |  E   O   F  |  F  W   Y   | U
   | 1 2         4|3 1 4        |             |
```

As you proceed, you may get a skeleton square looking something like this:

```
   1 2 3 4
1  Z I D E
2      X   Y
3  P       R Q T   etc.
4      F
     H O
```

(Note we can place QRT on the same row as P, but haven't yet assigned the column number. Similarly, H and O are placed by column, but not by rows.

When you can chain these results no further, you have several options. A good guess would help — and there is ample opportunity in the -EXPE–EDD.

Personally, I'd also be tempted to play around with the square, but at this point you will probably do better by putting in the obvious plain word.

Eventually you will have most letters assigned numbers, and will be able to fill in the cipher, which reads horizontally, and the plain text, which is formed from the vertical pairs.

```
                C K E X P E ....fill in plain as you get values
#1 |F   Q   C  |I   P   F  |Z   Z   S   R   V   D
   |4 3 3 4 4 5|1 2 3 1 4 3|1 1 1 1 etc|
   |           |           |           |
   |I   U   K  |F   E   O  |F   Q   T  |O   W   U
   |1 2 2 1 5 4|4 3 1 4    |4 3  etc.  |

#2 |F   Q   C   Z|X   E   P   D|Z   S   R   X|A   R
   |4 3 3 4     1|2   etc.     |             |
   |  Y   N   B  |  E   O   F  | F   W   Y   |  U
   |1 2 2 1 5 4 4|3 1 4        |             |
```

Of course it helps, when guessing, to take advantage of "maybes" from the square. For instance, if you DO put in EXPECTED. you will know the next word can't start with DZ or DD, so the plain is either DI, DE (or possibly 15, whatever that may be). And therefore Q must be 32, 34 or 35.

Given the T, we note that it is now the second plain letter. What is the first letter? In row 4 we have F and C. And in the proper column (1) we have the impossibles Z, P and H—but what about A? (Of course it could be some letter we haven't met with yet, S being a good candidate.) Let's try A for 41. (We can always erase the result, and try S instead.)

Now what might be "expected" that goes AT——? (It could be A TRAIN expected — but that would hardly get Zendia in the panic it seems to be in.)

Go on, finish it.

6 x 6 Bifid

A variation on the Bifid to which we were first introduced utilizes a 6 x 6 square, with all 26 letters and the ten digits inscribed within.

Since one has to juggle more coordinates, and there are on average slightly fewer letters per row, the job is a trifle more difficult than solution of a 5 x 5.

However, once you get used to it, there are certain features that should be of great assistance.

There is usually a rule governing the placement of the digits within the square. Number one must be next to the letter A — usually to the right of it, unless A comes at the end of a row (in which case the first thing on the next row is the digit 1).

However if the square is made up as a spiral, or any other rules of direction for inscription were followed, No. 1 follows A in the properly designated way.

Since J is the last letter having a number association (0, as the logical thing for the 10th letter) you know the last couple of rows of the ordinary keyword-mixed square

will probably contain six letters each, progressing towards the end of the alphabet. (Spirals, etc., excepted.)

The cryptographic function of the 6 x 6 naturally is more suited to traffic with dates, and other numbers. (In a problem you worked on above, you saw how much effort goes into writing 1938.)[4]

Here's a cipher — and you can see that it has both letters and numbers in the text, which is conveniently written out in 6-wide groups:

```
LDSTDY MGIF4B 9QAOZT 4T15JB IAONCM 3AOK7T OR71K4 MRDVAI
YBB9I0 BCR04P FL56BE D8AFBR OWND1W A64L45 ALETBE SAOKTR
IOOS80 CRF4P6 L64B45 RRUIKQ N8FZ78 OMIID7 PH
```

The text is evidently talking about "Satellite statistics" and two cribs are given, INACIRCULAR and SEALEVEL, both unplaced. Since we have stated our object to be "learning" rather than "solving" only, we'll give you the benefit of a beginning: AGEOSYNCHRONOUS.

The text on an even width (6 or 8, for instance) presents a somewhat different problem than a message of a width 7 or 9, and you will get used to taking advantage of the helps and swearing at the hindrances. (As the British say, what you lose on the swings, you gain on the roundabout.)

As in the Bifids you have just solved, the plain is read vertically, while the cipher is read horizontally.

```
AGEOSYNCHRONOUS
L D S M G I 9 Q A 4 T 1 I A O 3 A O O R 7 M R D

T D Y F 4 B O Z T 5 J B N C M K 7 T 1 K 4 V A I

Y B B B C R F L 5 D 8 A O W N A 6 4 A L E S A O

9 I 0 0 4 P 6 B E F B R D 1 W L 4 5 T B E K T R

I O O C R F L 6 4 R R U N 8 F O M I P

S 8 0 4 P 6 B 4 5 I K Q Z 7 8 I D 7 H
```

Because we have a crib at the beginning, we should take advantage of it, and build up as much of the square as we can. However, it is possible to place cribs on an even width, just as in the case of an odd width.

The reason is similar to that discussed in our "one-out-of-five" chance in earlier paragraphs. In the text above, there is a 1 in 6 chance of the first plain being an A, and a 1 in 6 chance that the second letter will be an L.

If you don't see why this is so, you might turn back to the 5 x 5 Bifid — and do some enciphering and deciphering on scratch paper.

This means that (in keeping with mathematical rules of when to add and when to multiply) there is roughly a one in 36 chance of having a vertical pair (such as the IN in the middle of the top two rows) represent the same plain IN. (Remember in the 36-letter alphabet there are 1296 digraphs, of which most are not very good candidates for plain text.)

Remember how "hits" between plain and cipher helped in the 5 x 5? Same thing here. And on an even width, the chances of a couple of digraphs in the crib being repeated in the cipher are in your favor.

I usually write down the running digraphs of the crib, and look for pairs in the cipher. But as a minimum, you should consider all "good" pairs as likely candidates for plain text, even if they are not in your "tip". For example, I see cases of IN, AR, AT and other juicy tidbits in our cipher.

Do any of them match with pairs within our crib?

Not only do we find a match, we find a consecutive occurrence of IN and AC — very nice for our crib INACIRCULAR. And one can hardly overlook the SY under SYNCHRONOUS.

This gives us so many goodies to work with that the other "tip" as well as obvious bits of text should pop out all over the place.

One of those many small points you might not remember now, but will ingrain in your mental computer if you do many 6 x 6 Bifids, is the fact that when a plaintext letter matches the cipher under it one position to the left, the plain value must be of an XX pattern — 11, 22 etc. Thus, when you get a value (either row or column) for O in the case below, you will get a two-for-one deal.

I happen not to have started with it, since I saw that A will also be an XX pattern, if you "take a few steps".

You should be able to start your numbering (with a letter that promises to give a lot of info) and build up the square and text rather easily.

We'll get you started:

```
AGEOSYNCHRONOUS AT       INACIRCULARO   T
L D S M G I 9 Q A 4 T 1 I A O 3 A O O R 7 M R D
14 5   1455 53  11 115  5 11551311555551
T D Y F 4 B O Z T 5 J B N C M K 7 T 1 K 4 V A I
15 5   2 15 55  15    5   12 125 115  25

Y B B B C R F L 5 D 8 A O W N A 6 4 A L E S A O

9 I 0 0 4 P 6 B E F B R D 1 W L 4 5 T B E K T R

I O O C R F L 6 4 R R U N 8 F O M I P

S 8 0 4 P 6 B 4 5 I K Q Z 7 8 I D 7 H
```

Solving the rest of the text is a trivial problem, but you might want to think about square reconstruction. You will want to take advantage of the fact that numbers "follow" the appropriate letter, and will note that 3 is on the same row as C.

Doesn't this imply, (on the relative numbering system I happened to start with) that the row and column sequences (which are the same) have 2 and 3 as consecutive numbers? And if you look further down, you find that 7 and G are both on Row 4 — making the sequence 5 1.

If you have pushed the plain along, you will have discovered that the digit 1 also fits on Row 1, and there is no place for it other than 16. So we must have A going to its proper number, and the sequence is 1 6.

Now we have 516423 or 516234 as the names for the row/columns. Let's look at both possible squares, with just the values we have recovered even before you finished the obvious text:

```
  5 1 6 4 2 3          5 1 6 2 3 4
5 O R     I 9  B     5 O R   I 9     B
1 T A 1 L C 3        1 T A 1 C 3 L
6                    6
4 G 7 h 8 j 0        2 K           F
2 K                  3 U
3 U                  4 G 7 h 8 j 0
```

The H and J (with attendant numbers) have to be placed on Row 4 as we have entered them, of course. And one must fit in B 2, and I 9 as consecutive values.

Now, which square would you choose? You of course realize that the arbitrary sequence beginning with 516 is actually equivalent to the 123456 the originator used; but we need not change the values unless we want to be extremely tidy, since it "works" either way. You will want to fill in any missing letter/number values, naturally.

A few comments pertaining to the solution of Bifids may not be amiss. One of the first problems may be the determination of the cycle in some cases (although this is usually given to you.)

In Twin Bifids the phenomenon of the "hit and miss" pattern of repeats is a dead give-away (in the rare event that you are not told what you are dealing with) but a related phenomenon may assist you in finding the cycle of a single message, as well.

If you have a significant repeat — say 5 letters — in the plain text, half of the time it will be on the same "cut". (I imagine by now you are picking up such terminology, and know that in a Bifid system a word like THERE could fit under the cipher letters as the digraphs TH ER EX or XT HE RE.

Incidentally, although I said "under the cipher", I find that many people invariably write the plain text for any given problem ABOVE the cipher values, while others enter the plain UNDER the cipher. It makes not the slightest difference as long as you know

what you're doing, but it sometimes poses a slight problem in a cryptogram appearing in a magazine, if the spacing between lines is not obvious.

Suppose we enciphered THEREISATIME, and later on THEREPORTER. One time out of two, we would get something like this, on a seven-wide cycle:

```
|T H E R E I S|A T I M E . .|      |T H E R E P O|R T E R     |
|p   b   d   i|a   p   b   u|      |p   b   b   q|c   b   x   v|
|1 4 3 1 3 2 2|2 1 2 4 3 . .| and  |1 4 3 1 3 1 4|1 1 3 1     |
|  y   l   x  |  z   e   a  |      |  y   l   o  |  w   s   g  |
|5 1 3 2 3 5 4|1 5 5 3 3 . .|      |5 1 3 2 3 4 4|2 5 3 2     |
```

with resultant text of

P B D I Y L X A P B U Z E A X etc., *and*
P B B Q Y L O C B X V W S G X.

I know it doesn't sound like much (and may be murder to find) but you will note that the digraphs which I have marked repeat with the same interval between the pairs.

Naturally a longer repeat would give a longer hit. In the case given both words (THERE) happen to start in the first position of the cycle of seven, but had one of them started in the third or fifth position we would have still seen something.

I do not want to get bogged down in a long explanation, so will leave the problem for you to think about, but you will readily determine that "broken repeats" of various kinds will show you a convenient entry into many phases of solution.

For one thing, you will see cases where you know that part of the repeat is on the cut while part is off the cut (whenever they start at different positions of the cycle.) These "same plain, different cipher" situations are invaluable (as you saw in the Twin Bifid.)

Once you have written text out on the proper cycle, you will find many repeats which cross the "border" of the cycle — some you are sure of, and some you only hope will pan out. Thus, when you see something like this, you can fly:

```
b d x y g|o r k u u|        |x y f d z|k u j g f|
                      and
 q t x b | t t c e |        | x b r d | c e o m |
```

Can we hope that the repeat extends over the demarcation line between elements of the cycle, and extends from "x y" through "k u" at least, and may extend in either direction for a few letters?

You could probably get part of a square by working with the above, but it wouldn't make any sense, since I just used the pattern to illustrate how it appears.

However, there are a couple of other minor points which some solvers seem not to take into consideration, and happily show up in the following example:

```
a r x o|w s f w|
 b c r | z f x |
```

It may be obvious that the second stretch of cipher ends with plain W—F–, but not all beginners realize the importance of finding a cipher letter on the bottom that repeats in the next position above, as did the "f" towards the end.

Even less "obvious" is the fact that when a letter repeats on the top level in the first and last positions of the cycle (as did the "w" above) the first plain letter of that cycle group is identical to this repeated cipher.

Finally, it may be worth taking a count of (or at least looking at) the cipher pairs obtained by taking each letter on the top row and indexing the next letter on the bottom row. Obviously the plain text at the positions of the bottom cipher under consideration are identical in case of such repeats.

Similarly, repeats of pairs counted from the bottom row of cipher to the next letter on the top level of cipher also have a plain repeat (at the position of the top level being counted.)

Remember that you can count the pair taken from the last position on the top row to the first position on the bottom row an appropriate number of letters preceding, when you are dealing with the first category above — from bottom to top.

(As in many of the "it is obvious" statements, if the matter seems a trifle obscure, a little thought and pencil-pushing will bring things into focus.)

C. M. Bifid

The CRYPTOGRAM has introduced still another (and more complex) Bifid into its pages. This is the C. M. (standing for Conjugated Matrix) Bifid.

The complexity (and attendant difficulty) comes from the fact that two entirely different squares are used — one for the plain text and another one employed for the encipherment (by recombining the resultant horizontal stream).

The increase in complexity is something like that of K-4 over K-3, or QUAGMIRE IV over QUAGMIRE III, since certain relationships are confined either to the plain or the cipher, and one must take care not to step over the boundary.

If one remembers to build up one square for cipher values, and one for plain values, the procedure is the same as you have just seen, so we'll give you a sample to solve.

Again, you will be given much more of a crib than is usually offered, since this first exposure should not be a challenge to your expertise, but an opportunity to assimilate techniques and procedures.

Titled "Sign is on again." Begins "IF THE UNITED STATES IS A MELTING POT THEN"

```
GGMDY BIMER DMMVD IIBWX ILGPM QSGUL GRLYF AYBBS LBIFR IMCTW
OLAWL EIZDW EIKDA IIRMH IOCVN TIHUB LTFXO NINDS TRKPH ILUOQ
GPVHS MTQU
```

It might be nuisance to determine the cycle, so we'll tell you it is five — coincidental to the fact that it is sent in fives. And only sufficient cipher is written out to cover the tip — when you finish it you can prepare your own cross-section version.

If there is any option as to a "good" letter to start with, there is some merit in choosing a good plain value, such as E, since when the time comes for testing, you are more likely to get words. (Alternatively you could choose a cipher that happens to be repeated many times, since it will undoubtedly be under a few good letters, just by chance.)

Call something 12, and see where it leads:

```
I F T H E U N I T E D S T A T E S I S A M E L T I N G P O T
G   G   M|B   I   M|D   M   M|I   I   B|I   L   G|Q   S   G
1   1   1|  2 1 1 1|  1 1   1|1 1 1 1  |1 1   1 1|2       1
  D   Y  |  E   R  |  V   D  |  W   X  |  P   M  |  U   L
    1   2|      1 2|    1   1|2        |  2   1  |          1
```

Cipher Square

	1	2
1	I R	G M
2		P
	D B	
	V Y	
	L R	

Plain Square

	1	2
1	T E	I S M
2		N

Do you see which is the best area of attack, and why? It is the Row 1 of the plain, since we have placed something in 11, which means we can proceed safely with the rest of the numbers; also there are five values on that row, which will take care of 1 to 5. (Cipher Row 1 has only four values, and the excellent plain values on Row 1 of the Plain Square should certainly lead someplace.)

So we'll put in the three letters for 3, 4 and 5. In doing so, you can try to prove you have that extra something that the TV ads are always talking about. It may not lead anyplace spectacular, but why not put the letters in some sensible order, just in case we are dealing with part of a key?

TEISM could be SMITE and ITEMS and MITES, as well as TIMES. Which would you choose? (Remember, this is not costing us anything — it's just a chance for a free "lucky draw.") I hope you chose TIMES, because if it is the top row, the 11 position of T looks great. (Do you see that if you had chosen MITES, the 11 would have made this the order on the THIRD row; and if you had chosen SMITE the 11 would have demanded that

this arrangement be on the FOURTH row, since the column and row numbers are the same.) We could start all over again, and name the columns and rows 1 2 3 4 5 — but that's just a point of nicety which does not affect our reading of the message, so we'll work with what we have put in already.

```
I F T H E U N I T E D S T A T E S I S A M E L T I N G P O T
G   G   M|B   I   M|D   M   M|I   I   B|I   L   G|Q   S   G
1 4 1 4 1|3 2 1 1 1|  1 1 3 1|1 1 1 1 3|1 1   1 1|2       1
  D   Y  |  E   R  |  V   D  |  W   X  |  P   M  |  U   L
3   1   2 3 4 3 1 2 3 5 1   1 2 5 3 5   4 2   1 3 4       1
```

```
Cipher Square          Plain Square
  1 3 4 2 5              1 3 4 2 5 1
1 I M G R              1 T I M E S
3     B                3   U     A
4   E                  4         F H
2         N P Q        2     N
5 V W     X            5
  D   Y
  L                        D
```

Again I stress that when Good Fortune lays her head on your lap, don't ignore her. An experienced solver would immediately cry "Eureka." For I'm sure you are pleased to find the letters bearing out the straightforward keyword mixed order, with no spirals or fancy work.

The implication is strong that U is in the plain keyword, and certainly the letters in the last three rows of both squares are "left-overs". Note that D COULD come before E and Y COULD be the next to the last letter in the first square.

Try the Y as the most likely bet, and see what it does. It immediately tells us that H is 45, a very likely spot, making F plain 44 or 42. This pins down the D cipher ahead of the E on row 4 — and so we roll. I'm sure you need no further suggestions, and will get the squares quickly. Be prepared, however, to spend a little more time on the next C. M. Bifid you tackle. The entry points were dished out very liberally in this example, and when we come to the point of discussing "What do you do when you don't have a crib?" you'll put this cipher in the Difficult-If-Not-Impossible category. (A quick answer to the question posed, is that you often have to MAKE your own crib. And it certainly helps to have electronic assistance.)

NOTES ON CHAPTER VI

(1) Later on we will discuss the recovery of squares in detail. For the time being, you should be aware of the fact that there are squares of the type shown above, and also several variants on the method of inscription.

Looking at the following examples, you may be interested in determining for yourself what some of the "rules" might be:

```
S W E T H    L O V E R    E D S A B    R L D M V
M N O P A    K A B M C    T Q R V C    E I F U W
L Y Z Q R    T D F G H    I P Z W F    P C G Q X
K X V U B    S Q P N I    N O Y X G    U A H S T
I G F D C    U W X Y Z    U M L K H    B N K T Z

D B F U V    U W X Y Z    B D E F G    C Y O B I
E I G S W    T M S V I    A V W X H    R T M H Q
M T H Q X    Q E P R N    U R Z Y K    P A F N V
O A K P Y    K L B O G    I Q P M L    G E L U X
C R L N Z    H F D C A    T S N O C    D K S W Z
```

The keyword (or key phrase) and method of inscription are rather easily determined in the pristine state of the above, but if you had a relative solution in your partial recovery would you recognize this:

```
  F   A   P
  U E X G
  H T Q
S W K Z
O       C
```

Of course such a relative solution works just as well to decipher the message, but solvers usually like to reconstruct all the elements, to make a "neat" package. And some squares give identical results in several forms — for it really doesn't matter whether you start at the bottom and read backwards or start at the top and read forwards, since the 180-degree turn makes the square and keyword look just as logical.

(2) Terminology may seem a trifle loose in some of the esoteric fields we are discussing, and it may not help to look up certain words in a dictionary. We are used to calling two letters a "digraph" and two numbers a "dinome" — but a dictionary may leave you under the impression that you must have two vowels, or possibly two consonants together to qualify as a "digraph", and "dinome" is probably not even listed.

So don't be alarmed if you hear the word "digram", or have to guess what "bisecting a message" means, or you find that "fractionation" is a chemical process. We will try to be precise, but remember that some of the techniques were developed — and named — by analysts, not semanticists. We know what Friedman meant by "indirect symmetry", but you probably won't find it in your college dictionary.

(3) An interesting practical exercise in cryptanalysis was developed by my good friend Lambros Demetrios Callimahos, who died a number of years ago. During the last phase of his government career he was a pedagogue par excellence. He conducted an ongoing course for middle-level analysts in which they were exposed to many classical systems and methods of attack.

One of the problems tackled consisted of a complete package of communications and background on a mythical country — Zendia. The fertile imagination of Lambros had dreamed up a government complete with dictator and its attendant problems of transportation, economics, etc., and the specific running of the military machine presently involved with aggression of some sort or another.

Codes and ciphers of many types were represented, including machines, and students gained a very good insight into the task of attacking a wide range of interrelated problems.

Even "outsiders" now consider the story of the Dundee Society, and more particularly the legend of Zendia, as a classic. The name "Dundee" was derived from the fancy marmalade jar sitting on the Teacher's desk, wherein his meticulously sharpened pencils resided.

(4) Some literal systems use a simple device for numbers. With some symbol to indicate digits, or by a convention such as enclosing a string within ZZ–ZZ, the letters A to J may conveniently express figures. Thus ZZAICHZZ is a lot more economical of space than the spelling out of the string of such as "ONE NINE THREE EIGHT."

Chapter VII

OTHER FRACTIONATION SYSTEMS

Digraphid

THERE are a number of rather interesting digraphic systems, some of which deserve specific attention here. One of the most interesting to my mind is the Digraphid (which name hints at a digraphic system having certain characteristics of the bifid systems already discussed.)

This system employs a horizontal box or table or rectangle, whatever you choose to call it, (9 wide and 3 deep, with digital column names, no zero) and a vertical box (9 rows of three letters each, with the keyword starting down the first column, and with the rows numbered in 1-9 sequence).

The boxes are usually based on a keyword, and the sign # makes up the 27th value, either as the last entry or following right after the keyword. These rectangles are connected by a 3 x 3 numerical square (in which the numbers 1 to 9 appear in numerical order) placed at the right of the horizontal table and directly above the vertical table. For example:

```
1 2 3 4 5 6 7 8 9
P R O B L E M A C   1 2 3
D F G H I J K N Q   4 5 6
S T U V W X Y Z #   7 8 9

                    1 S C Q
                    2 O D R
                    3 L E T
                    4 V F U
                    5 I H W
                    6 N J X
                    7 G K Y
                    8 A M Z
                    9 B P #
```

The final digraphic cipher is composed by taking (as an intermediate step) a trigraphic substitution of the plain digraph by reading the column name of the first letter and the row name of the final plain value, and inserting the digit appearing at the intersection in the 3 x 3 box. Thus TH would be temporarily replaced by 285, using the above components. The string of trigraphs so obtained is written out vertically, one plain value after another — 285, 622, etc:

```
2 6 3 8 4
8 2 6 2 4
5 2 2 9 5 etc.
```

You can see the plain that I used, if you realize that 285 represents TH and the next 3 digits (622) represent ED and so on. This intermediate form should not be confused with the cipher, for the next step involves taking the numbers out of the box HORIZONTALLY (remembering that the plain text reads down the columns VERTICALLY.) This is done by reading from our tableau (as some table arrangements are called) the equivalent to 263, and so on, sticking to some preindicated width. Suppose we are dealing with a four-wide agreement. Then the second cipher would be 882, the next 625 and the fourth 295 — which turn out to be cipher letters (from the tableau) FT, ZD, EH and RP. Sent in five-letter groups, the cipher would be FTZDE HRP..... and so on, but the constructor may be kind enough to break up the digraphs into clusters on the correct width.

So much for encipherment. To grab a toehold on the solving process, the simplest method is for us to solve (together) a problem.

```
ILOSSRCULE RPRQTHWPHN BQAVMEWXSB EMJJANF#AM UWRAQBEYOK
                                  turnsonhis leftturnsi

GNLOIPBL#T CMABDPNEWS CBUDNQRNOY WQIZ#BGZEF QVMMBITEIY
gnalonlywh enhe

RPJQONJSRA KKEL#LSOSO EPFDBK
```

You will note a crib has been placed for you. Since this amount of help requires a fairly experienced solver, and we want this present exercise to be a demonstration of "how to go about it" rather than a "Gee, ain't I clever." dissertation, I have added a bit more of a clue by giving you the beginning "I would pass," which might lead a really adventurous cryppie to the next portion — "a law that any driver who." (There are technical reasons why this would be a good guess, such as the "RP RQ" cipher which fits conveniently under the plain "AlAwthAtAn" for reasons you may see later.)

The cipher text is written into a work sheet providing enough room to assign "values" as the tableau is recovered. Think of your work sheet as being a skeleton version of the illustration I gave above, showing how the system works. The plain text I have written in small letters above the cipher, in the proper columns.

```
iw ou ld pa ss al aw th at an yd ri ve rw ho tu rn so nh is
IL       OS   |RP       RQ   |BQ       AV   |EM       JJ
              |               |               |
   SR       CU|   TH       WP|   ME       WX|   AN       F#
              |               |               |
      LE      |       HN      |       SB      |       AM
```

```
le ft tu rn si gn al on ly wh en he
UW        RA    |GN        LO    |CM        AB    |CB        UD
                  |                |                |
   QB        EY|   IP        BL|   DP        NE|   NQ        RN
                  |                |                |
      OK        |       #T       |       WS       |       OY
                  |                |                |

WQ        IZ    |QV        MM    |RP        JQ    |KK        EL
                  |                |                |
  #B        GZ |   BI        TE|   ON        JS|   #L        SO
                  |                |                |
     EF         |       IY       |       RA       |       SO

EP        |
          |
FD        |
          |
BK        |
```

You will note that the final segment has only three digraphs, which necessarily fall into the pattern shown.

In many system solutions, it pays to organize your worksheet with such an area as "what you know" about certain components. In this case, before jumping into suppositions or blindly searching for a likely place to start, it is wise to make a listing of values in the first (horizontal) tableau that are in the same column, and values in the second (vertical) tableau that are in the same row.

Referring to the text and crib above, it should be obvious that P is in the same column as O in the first tableau, since any value you assume would start a cipher text numerical value for OS (reading horizontally) with the same digit as the plain value PA (reading vertically).

Even more obviously, the first cipher value for IL has to start with the same digit as the plain value IW. Which gives us an opportunity to underline a great weakness of the system, which permits of crib placement quite frequently.

A little reflection will make you realize that there is one chance out of three that a given value for a plain digraph will result in a cipher value that starts with the same letter. After all, there are only three letters in each column of the first tableau. Similarly, you will find that the last plain digraph in any cycle segment (five wide in the above message) ends on the same digit as the final cipher digraph of that segment of the message. In the case at hand, the relationship of plain SI to the cipher OK says that I and K are in the same row of the second tableau.

Without belaboring the point, you will find that there are several situations, at different cycle widths, which permit of this "in the same column/row" classification. For instance, look at the LD plain digraph and the SR cipher in the first block above. Can you see that the letter L must be on the same row as S in the first table, and D must

be in the same 9-deep column as R in the last table? No matter what arbitrary value you assign to the intersection of this "cross" of SR and PA this will be true.

Let's look at a skeletonized version of our cipher with the proper plain digraphs put in, and see what relationships we can determine. (I have started to make a few arbitrary number assignments, which will be explained below.)

```
iw ou ld pa ss al aw th at an yd ri ve rw ho tu rn so nh is
IL       OS    |RP       RQ    |BQ       AV    |EM       JJ
   3     3     |      1        |                 |1
   SR       CU |   TH      WP  |   ME       WX   |   AN       F#
            1  |   1           |                 |2
      LE       |      HN       |      SB         |      AM
   3           |               |                 |3

le ft tu rn si gn al on ly wh en he
UW       RA    |GN       LO    |CM       AB     |
      1        |      3        |1                |
   QB       EY |   IP      BL  |   DP       NE   |
      2     1  |               |                 |
      OK       |      #T       |      WS         |
      3        |               |                 |
```

First, let's write down all column relationship for Table I and row relationship for Table II — noting, for example that P is in the same column as O in the first box (or table — whatever), and so on. (You will naturally combine the I-E and E-C relationships into a three-deep set. Similarly ESM and KTH are row sets in the second box.)

Remember to consider such relationships as the cipher F# and the plain RN in the fourth segment of the cycle — for a little thought will enable you to write down # and N as being in the same row of Box II.

Table I, in same Column: P O, T E C, N J, A R, Y B and L U.

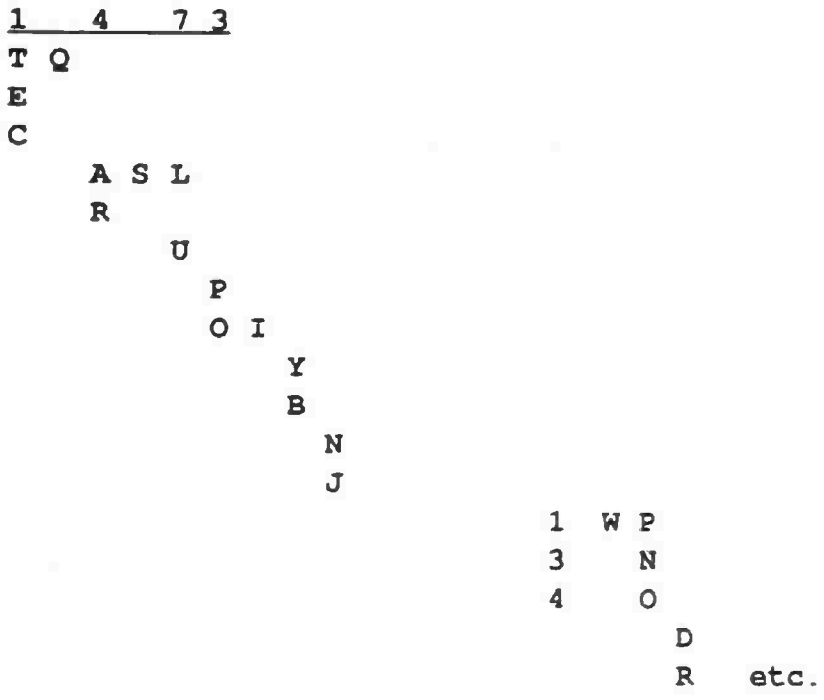
Table I, in same Row: S L A, Q T, I O, and MV

Table II, in same Row: E S M, B O, K I X, T H Y, P W and # N

Table II, in same Column: R D, N O P and B U

Make sure you understand why the above relationships are valid.

We can begin to make a crude attempt at the tableau reconstruction by compressing the above values to some such arrangement as the following, which will serve as a worksheet for assigning arbitrary values. (I will give the columns in the first box and the rows in the second box numbers, being careful not to use the same number twice unless and until the tie-in is proven. I will also be parsimonious in my assignments, making sure I choose the letters first which will have the most effect on other letters.)



As values become apparent, I will put the correct numbers in the (ultimately) 3 x 3 area to the right of the first box and above the second table.

Note that if I arbitrarily call the TEC column 1 in the cipher write-out above, then it might be a good idea to call TU 123, using the next digits available, since TU appears twice in the text. (I will then soon try to call RN 456, for the same reason, but quickly find that the first and second digits must be the same. so settle for 445. Remember that the 5 could be 3, which has just one letter in a row, but that will come out in the wash even if we assume a new value, 5.)

And note that the cipher JJ becomes ..2, so that I have put J on row 2 of the second table.

So things start to snowball, with such discoveries as W has to be 1 on the second table, CU has to be 113, and so it goes. In the following write-out, I have continued assigning numbers, remembering not to use a previously assigned number until a relationship is proven.

One other important point. Inspiration is supposed to be the better part of solution, in some cases, and without jumping into the probable text you would have had a murderous time, and possibly reached a dead end. (We can always backtrack if the guess turns out to be a lousy one.)

```
|le ft tu rn si gn al on ly wh en he
|UW       RA    |GN       LO    |CM       AB    |
|7     1  4  6  |   4  3  7     |1              |
|   QB       EY|   IP        BL|   DP       NE|
|      2  4  1  |4  9  4  1     |4              |
|      OK       |      #T       |         WS    |
|      3  3  8  |   7  5        |3              |
```

You will soon get enough values in your skeletonized tableaux to see that M N O P Q are probably all in the same column of the second table (and probably consecutive.) When you consider the letters that belong on the same row, and those in the same column, you rather quickly come up with:

```
6  S  E  M
3  U  #  N
4  B     O
1  W     P
         Q
```

This must be the correct order of columns, since the location of the # implies a 10-letter keyword if the "SUBW" begins it, which seems likely; and the M-Q is necessarily in the last column because there are not enough letters from R to Z to fill a complete "last column."

Obviously 6 3 4 1 9 should be 1 2 3 4 5, but don't worry about our equivalent solution — it is just as valid. Since the ACA uses the normal order for all boxes, including the 3 x 3 intersecting one, you can take advantage of this in many ways, such as hoping that # will be in the 9 column (unless it follows the key word) and knowing that it is unlikely for low letters of the alphabet to be in the last row of the first table or the last column of the final rectangle.

Remember that if you have decided that 6 3 4 1 9 should be 1 2 3 4 5, etc., as in the above paragraph, you must use the recovered relative order in the 3 x 3 connecting box as well—unless you wish to take the time and trouble to put everything in pristine condition.

I usually wouldn't bother to change relative-to-true, but occasionally I find it possible in the early stages to decide that such-and-such a letter combination should be on the bottom row of square I and the third column of square II. In such cases, I would take advantage of the fact that the middle digit of the cipher trio should be 9. Similarly, in case you have a J (which is probably in the middle column or row) I might use such information in starting the arbitrary namings. One always lives in hopes that the first supposition may be the right one, and there is nothing wrong with giving Fate a nudge or two.

It would make the system more realistically difficult to solve if the numerical sequences were scrambled — possibly three different mixed sequences to recover.

But since this type of solution is supposed to be a fun exercise, let's let well enough alone.

Since that order of digits represents the names of the columns in the first box, let's look at our recoveries there:

```
6 3 4 1 9
  O R E I
S   A
      T
      C
```

Doesn't it seem likely that the T is on the last row, since it has Q on that row — which places the C next to A, if that begins the "left-over" portion of the alphabet, and B must be in the keyword. The -OREI— certainly suggests FOREIGN, and the ultimate combination has to be FOREIGNBUS.

We fortunately had enough of a crib to make the solution relatively straightforward, but the problem can be extremely tough if little is known.

The chances of plain/cipher values being identical in the prescribed places helps a little in guessing text, but you would have had to take an incomplete tableau and apply it to the rest of the message, trying to break into the text at other places, if you had not been given so much.

And especially keep in mind what I said about "inspiration" above. My wife kept on running into snags during the complicated phase of assigning digits, one after another, and finally took advantage of the M N O P Q order in the last column of the final box.

Using the arbitrary values for these 5 letters in all three places (first table, and 3 x 3 box as well as the last table) things fell to pieces immediately.

Foursquare

I personally find this system not one of my favorites, since it involves the chore of testing the crib in two possible positions, and looking for a matching cipher relationship, after which the solution is largely mechanical. But one man's meat, etc. — so here goes.

As the book says, four 5 x 5 squares are set up, the first and third being normal order (no J of course) but the second and fourth consisting of different mixed sequences:

```
A B C D E   G R D L U
F G H I K   E Y F N V
L M N O P   O A H P W
Q R S T U   M B I Q X
V W X Y Z   T C K S Z

L I C N V   A B C D E
O T D P W   F G H I K
G H E Q X   L M N O P
A M F S Y   Q R S T U
R B K U Z   V W X Y Z
```

The first letter of the plain digraph is found in square 1 and the second letter in square 4 (the lower right sequence.) The two cells thus delineated are considered opposite corners of a rectangle; the cipher digraph is found by reading the remaining corners of that rectangle, from square 2 to square 3. Thus plain TH becomes IP.

To place a crib, if there is a repeated digraph (or more) at an even width in it, the matter should be one of simply looking for an equivalent repeat in the cipher text. If there is no such favorable case, one has to encipher the crib in the way described above, and look for a place in the cipher text that has the specific pattern of isomorphism — repeated letters in the cipher at identical spots to any repeated letters in the artificial text one has obtained by enciphering.

One can do this either on some available matrix for any Foursquare problem, such as in the ACA handbook, or use the normal squares of 1 and 4 for plain text entry, recording the positions in squares 2 and 3 with numbers representing the sequent positions 1, 2, 3, etc.

Thus if the plain is "co me qu ic kl yw en ee dh el px" one could encipher the message on the square above as LE WI XA FN EX CU DX UV DP GX HZ. Note that in the first positions, the Dx and Dp would appear at the proper cipher position as a repeat, and similarly there are three X's in the second positions of digraphs .X .. .X .. .. .X. (Remember it is not the PLAIN version of the crib which one is trying to match by isomorphism, but the arbitrary ENCIPHERMENT.)

If one has merely jotted down numbers in the proper places of boxes 2 and 3, or writes the numerical pattern of the arbitrary encipherment of the crib, repeating numbers ONLY when the repeat is on the proper cut of beginning and ending letters (since there is no necessary correlation between an A in square 2 and an A in square 3) one would obtain a pattern of 1 2 3 4 5 6 7 8 9 10 11 12 13 10 14 15 13 16 17 10 18 19.

Suppose the crib happens to come on the alternate cut, appearing as ".c om eq ui ck ly we ne ed he lp x". The arbitrary encipherment would then yield a pattern of 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 14 13 18.

One gets an entirely different pattern to look for, since the only repeats are the 14 in the second position and the 13 in the first.

Once one has placed a crib, the recovery of the two cipher squares is a matter of filling in whatever letters are obtained by locating the plain letters in the first and fourth squares, and entering the first and second members of the cipher digraph into squares 2 and 3, respectively, at the positions marked by the upper right and lower left corners of the rectangle described above.

Once all the information obtainable from the crib is entered, the job is one of completing the cipher alphabets in the squares, by partially deciphering as much as you can of the rest of the text, making judicious guesses as to the plain, and figuring out what the keywords might be, and the order of inscription.

Here is a Foursquare problem we will work on. The crib supplied reads "IDECIDEDTOLOOK." Notice that the repeat of IDE means you must find a repeated digraph two positions removed, no matter which of the two possible cuts is correct. The SK ES SK position is the obvious choice, and the two possibilities are explored.

I usually would try them both at once by writing in the cipher values either in two colors or in upper and lower positions, trusting that one or the other would hit on a confirmation — but the presence of a convenient final letter repeat in the cipher of the ES .. RS makes the test very simple.

You will find that if you try plain EC for cipher ES and ED for RS you get the proper repeat in the second position, but that moving the crib so that you encipher CI and DT (instead of EC and ED) gives a repeat in the first position of the cipher digraph instead of the second. So let us work with the text in which the crib is properly placed.

```
HA UV RS VR HA WK QV OS WA MP YU VR UY XX YC ZR KC VR OK GK

VW HA KL OK BK PG UF VZ SK ES SK RS VR UI IQ RG EG LK ZS OI
                        id ec id ed to lo ok

IT RC EW CK VA RF EI OQ IU UL WQ RG NC EG RT RF LT CK
```

The first step is of course the filling in of the cipher into the proper position of the two squares:

```
A B C D E        E R
F G H I K        Q S
L M N O P      y z U I
Q R S T U      x W V
V W X Y Z

      K S    A B C D E
      Q      F G H I K
I     R      L M N O P
                Q R S T U
                V W X Y Z
```

If you have had any experience with alphabet square construction, you will immediately recognize the method of inscription.

Although the upper right square has R S U in the fourth column, and almost seems constructed by writing the keyword down successive columns, the presence of I in the last column knocks the idea out, and suggests that both squares were made by inscribing a keyword alphabet in spiral fashion, beginning ..ER..I in one, and ...KS in the other. (In any case, there are not enough letters after the RSU in the alphabet to take up the last column.)

In case your knowledge of square construction is sketchy, the term "spiral" implies that the keyword alphabet starts in the upper left-hand corner and progresses around in Right-Down-Left-Up (etc.) fashion, ending with the last letter of the sequence (Z often) in the middle cell. Of course you could go down, right, up and left —and start in any corner, but the results are usually equivalent.

You can also start in the middle and work your way outward, which gives quite a different square than the above.

If indeed the top square uses the spiral inscription, you can complete the sequence after U with VWXYZ — and try it on the dog, as they say. The lower square has three possibilities for each cell after R –TUV (you will note that S is used already) UVW, VWX, WXY and XYZ being the possible choices. Similarly, following I in this square, one would expect the sequence to run LMN, MNO, NOP before the Q, with three letters to choose from.

Deciphering what we can, we have the following, and I call your attention to the repeat of HA which may aid you in guessing plain:

```
HA UV RS VR HA WK QV OS WA MP YU VR UY XX YC ZR ......
   ed to    tc              or to mo rr l. on
ot?             is?             nr    no rm m.
```

Is there any doubt as to the proper choice of text?

Tomorrow stands out like the proverbially painful thumb. Putting this in, and entering the proper values in the squares, we have:

```
HA UV RS VR HA WK QV OS WA MP YU VR UY XX YC ZR KC VR OK ..
   ot?ed to    tc is? * *     ?r to mo rr ow on  * to *
                      ea qc                       dv    da
                      kb rh                       iw    ib
                      pc sn                       ox    oc
                      ud tx                       ty    td
                      ze us                       yz    ye
```

*Is it clear that when you have one of the letters of the cipher digraph on a square (but not both) you can limit the plain to five values for the first position and five for the last, giving you 25 possibilities for the total plain digraph?

```
VW HA KL OK BK PG UF VZ SK ES SK RS VR UI IQ RG EG LK ZS OI
rs       ** **    ** so id ec id ed to lo ok ad*
ss                                          bi
                                            co
                                            dt
                                            ey

IT RC EW CK VA RF EI OQ IU ULWQ RG NC EG RT RF LT CK

        A B C D E       E R
        F G H I K       Q S   (Note Q added since U placed)
        L M N O P     Y Z U I
        Q R S T U     X W V
        V W X Y Z

              K S   A B C D E
              Q     F G H I K
        I Y Z R     L M N O P
                    Q R S T U
              C     V W X Y Z
```

Following "decided to look" you get the possibility of "at" for the next digraph, which should make the rest of the solution simple, and I leave the completion of the squares up to you.

Two-Square

Because, unlike the Foursquare, you are not given any knowledge of the squares, this system presents more of a challenge in placing a crib and in the solution of the plain text.

Two 5 x 5 key squares are set up, for example:

```
D I A L O   B I O G R
G U E B C   A P H Y C
F H K M N   D E F K L
P Q R S T   M N Q S T
V W X Y Z   U V W X Z
```

To encipher plain text, it is divided into pairs; then the first letter is located in the first box and the second letter in the second box. The cipher equivalents are those letters forming the opposite corners of a rectangle with the plain-text pair. Thus, TH becomes QC. If (as happens one-fifth of the time) the second letter is on the same row as the first letter, the values are reversed for the cipher, so that DR becomes RD.

This reversal (when on the same row) helps place a crib — and if there is no reversal, or repeated digraph, crib-placement may be a matter of try, try, try. An illustrative problem is:

```
IA IF BO PH VQ PQ UW OF SZ ID NE GH ZH FO NE HE PT ZY GA

FM TE WL YF PI HA VH FN TE XP OK TH VE ZO IH WH WV MQ NO

EE OK TH AA FM TE WW ZH FZ FD HH
```

The crib given is "the handle of the big" — and it doesn't take much imagination to guess that the next word is probably "dipper."

I usually scan the cipher text looking for good reversals, such as HT, or ER (or RE) — any interesting combination might be worth circling — but in the back of my memory when so doing is the given crib, with the always present hope of finding a reversal contained therein.

Sure enough, there is HE in the cipher which might be EH; and, of all things, the PI comes at the right spot for its occurrence in the supposed "dipper."

Let's see how much of the two squares we can reconstruct on the basis of the crib placement.

```
IA IF BO PH VQ PQ UW OF SZ ID NE GH ZH FO NE HE PT ZY GA
                              th          th eh an dl eo

FM TE WL YF PI HA VH FN TE XP OK TH VE ZO IH WH WV MQ NO
ft he bi gd ip pe r

EE OK TH AA FM TE WW ZH FZ FD HH
           ft he
```

One merely forms the prescribed rectangles of plain and cipher values to reconstruct the square, hoping that columns and rows will be tied in sufficiently to permit further decipherment, or even a stab at the keywords.

```
          T  |  N
          E  |  H   G E
I         A  |  P   O
      D      |    Z
      Y      |    L
    F        |        F   D
    M   H    |        T
  B          |            W
  L          |            I
    G        |              Y
        r?   |                v?
```

Whoops. Someone must have made a mistake — either I or the person who created the problem. After putting in the I on the same row as P, to create the cipher PI, the next digraph should be PE plain, followed by R-something. But cipher HA is already determined as being EP on our tentative squares. (It's very easy to have such a lapse as enciphering EP when you mean to encipher PE.)

Not to worry. Mistakes do occur, no matter how carefully the editors check text, so let's do some deciphering and see what we can straighten out.

```
IA IF BO PH VQ PQ UW OF SZ ID NE GH ZH FO NE HE PT ZY GA
         a                    th e        th eh an dl eo

FM TE WL YF PI HA VH FN TE XP OK TH VE ZO IH WH WV MQ NO
ft he bi gd ip pe?r     he    ig?ht r?                 on
               ep?

EE OK TH AA FM TE WW ZH FZ FD HH
ee ig?ht ho?ft he            e
```

Even making a few guesses of plain (the "eighth of" looks good when you consider the slight confirmation of "on" plain being "no") doesn't get us very far. But looking at the squares, one gets the impression that the second one was constructed by taking a keyword and putting the rest of the alphabet in, following the logical order. Note that E G H and O P are on appropriate rows, and moreover, the order G H fits with the O P when you start a square:

When you start to play around with this idea, the first square looks more and more like a spiral — starting along the top row, down the right side, left along the bottom, etc.

The right-hand square seems to be a row-by-row keyword-mixed affair, and if you put WXYZ on the bottom row, OP on the 4th row, and GH on the third row you can begin to get the feel of both squares.

You can afford to try all sorts of possibilities if you keep some sort of record of your guesses. In the above, I merely put in a few possible positions with question marks, but you might want to use colored pencils, or write out the squares in different ways, etc. You will find that out of the hodgepodge something starts you on the right track.

Although it doesn't sound very specific, when you have one of the letters of a digraph placed, you often can limit the total plain digraphs to 25 possibilities, since you have the coordinates of a row on one square and a column on the other.

Note the beginning cipher IA. On the squares above, the first plain would have to be in the same column as A in the first square, and on the fourth row of the other. We have an indication of T E A and a possible H going to O P Q and possibly T. What digraph suggests itself? Maybe AT, but when you play around with the possibilities you discover a much better guess would be TO FI.... and perhaps "to find."

It doesn't hurt to take a flyer, as long as you know where you are and what you're doing — and can pick up the pieces if things break down. Putting in "to find", we get

more information on our squares and text. I have put in a few assumptions, and not bothered to straighten out the confusions — some of which may arise due to errors of inscription. See if you can recover the keywords.

```
IA IF BO PH VQ PQ UW OF SZ ID NE GH ZH FO NE HE PT ZY GA
to fi nd a     ur    ??    ow th ea    of th eh an dl eo

FM TE WL YF PI HA VH FN TE XP OK TH VE ZO IH WH WV MQ NO
ft he bi gd ip ep?r     he br?ig ht r? dd ta  a bo?ut?on

EE OK TH AA FM TE WW ZH FZ FD HH
ee ig?ht ho?ft he ci?rc le ov?er
```

Fractionated Morse

Fractionation is not limited to digraphs — as witnessed by the Fractionated Morse system, in which the dots and dashes of an original Morse plain are cut up in some manner and reassembled, or indirectly made part of a substitution process, to create a cipher version.

If you don't already know the Morse code, you may want to refer to the ACA handbook, or a dictionary, or a *Boy Scout Handbook*, or something else to tell you that A is dot-dash (.–), B is dash-dot-dot-dot (–...), etc.

The plain text is put into Morse, with the dots and dashes having a third element added — an x — which is placed between the elements of each letter. Two xes are used to separate words. The message "come at once" would be written as:

```
–.–.x–––x––x.xx.–x–xx–––x–.x–.–.x.
```

This series of symbols is then taken off in threes, and each three-element changed to a letter by referring to a prearranged table, in which 26 or the 27 possible combinations of dot, dash and "x" are assigned a literal value.

The elements (by convention) are written out in a progressive manner and the 26 letters of some keyword alphabet is written above, as follows:

```
R O U N D T A B L E C F G H I J K M P Q S V W X Y Z
. . . . . . . . . . – – – – – – – – – x x x x x x x x
. . . – – – x x x . . . – – – x x x . . . – – – x x
. – x . – x . – x . – x . – x . – x . – x . – x . –
```

Now the Morse equivalent which you prepared for the plain text is taken off in sets of three symbols, and the corresponding letter of the keyed sequence above the table is substituted. Thus, the first three symbols of your plain version being –.– the letter C is substituted, to become the first cipher value. The second set of three being .x– you must substitute the letter B, and so on. Your final cipher should be CBIIL IMHVV FL (the cipher text usually being sent in groups of five letters.)

A Fractionated Morse problem appears as follows, with the given crib "don't know."

O M O K Y O F Z H T C A K Y Q P A A Q S J G P A T E V A P L

Y O F Q O N X O F U P S I P O L S C J G N Z O F U X U J L G

I Z V J G V P J W P O Q U O F P J Z P O S I A D F U O W A V

Z C P J G H Y W Q L G G V Q Q O C L G W L M N L G W U O A O

The first problem will be to place the crib. Since we don't know where the crib starts, and there are three possibilities for the correct cut of three symbols, the placement could be:

xx- ..x --- x-. etc., or x-. .x- --x -.x or -.. x-- -x- .x-

— we must check for an isomorphic hit on any one of the three "maybes." The simplest way to do this is to write the crib out in Morse on three levels (starting with the word spacer xx) and assign values of a word pattern — A for the first, B if the next value is different, or repeating A if it is the same, etc.:

x . - x x x - . - . x
x . - - - - x x - - x
- x - . x . - - x -
A B C D E D F G H I

This isn't a very great pattern to find, since it has so few repetitions, but perhaps one of the next two will be more encouraging. (Of course we have to try all three cases anyway, but if there is an easier one to find — or eliminate — we might start with it on our final search.)

The second possible "cut" would be if the first value is x-. and it can be easily prepared by copying the second and third row above, and then writing the top row one position removed:

x . - - - - x x - - x
- x - . x . - - x -
x . - x x x - . - . x
A B C D E F A G H C

This is slightly better, with two repeats. The third should be:

- x - . x . - - x -
x . - x x x - . - . x
x . - - - - x x - - x
A B C D E F G H I J

which is the worst of all, with no repeats. (Of course, even that is a "findable" pattern, in that a stretch of ten unique letters must fit, although there will probably be many such stretches.)

Let's look at our cipher, and see how many places we must try for the first or second possibility. We must note cases of X–X and Y–Z– – –Y– –Z

I don't see anything except the UXU in the second line, so let's try our crib there, and hope it works.

```
O N X O F U P S I P O L S C J G N Z O F U X U J L G I Z V J G V P
                                  x . - x x x - . - . x
                                  x . - - - - x x - - x
                                  - x - . x . - - x -
```

You will note that the Z checks in the two elements we have — and what is more significant to the experienced cryppie, the XX- for Z is exactly what we would expect in a keyword alphabet if Z is the 26th letter.

We should now prepare two records, one for the symbols to place under a normal alphabet, so that we can easily find our known values, and another chart listing the logical order of progression through the dots, dashes and x's, so that we can spot a keyword when and if it appears, and keep an eye out for nice-looking stretches in our placement of letters.

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
          - -   . -   .     .           x     x   x
          - -   - x   x     .           -     -   x
          - x   - -   -     x           .     x   -

. . . . . . . . . - - - - - - - - - x x x x x x x x
. . . - - - x x x . . . - - - x x x . . . - - - x x
. - x . - x . - x . - x . - x . - x . - x . - x . -
    O   I     L         F G   J       U   X   Z
```

Isn't that a pretty picture? The letters in the "leftovers" fall into place in a way that adds a few values, and makes us very confident that we are on the right track. So we fill in the H and Y of our equivalents, and then apply what cipher values we know to the text.

```
O M O K Y O F Z H T C A K Y Q P A A Q S J G P A T E V A P L Y O F
.   .   x . - x         x x x     x   - - x           x . x . -
.   .   x . - x         x             x -               x x . -
x   x   . x - -         .             - x               - . x -
I       /S O            /               O                 T /S O

Q O N X O F U P S I P O L S C J G N Z O F U X U J L G I Z V J G V
x .   x . - x x   . x . .     - -   x . - x x x - . - . x   - -
  .   - . x - -   -   . x     x -   x . - - - - x x - - x   x -
  x   x x - .     -   x -     - x   - x - . x . - - x x -   - x
  l u   T I O N         E       y O u   D O N T / K N O W /     O
```

```
P J W P O Q U O F P J Z P O S I A D F U O W A V Z C P J G E Y W Q
x -   x . x x . - x - x x .   .     - x .       x   x - -   x   x
  x   .     - . -   x x   .   -     - - .       x     x -   x
  -   x     . x -   - -   x   -     - . x       -     - x   .
        /     B O     T / T           O B         /     T O     / £
x

L G G V Q Q O C L G W L M N L G W U O A O
. - -   x x .   . -   .     . -   x .   .
x - -       .   x -   x     x -   - .   .
- x x       x   - x   -     - x   . x   x
O M             O           O     B
```

Examining the text, one finds quite a bit to push on with. The first obvious data comes from the necessity for both P and Q to start with "x" since at the end of the first line you have a plain of S followed by three dots — a guaranteed SO since there is no four-element letter starting with three dashes.

Similarly, P is seen to start with an "x" and I have filled all P and Q occurrences in accordingly.

Now comes the chance to make some logical guesses. If you agree that the suggestions made (using lower letters rather than capitals) are sensible, you should have no problem filling in both the text and the alphabets.

Trifid

Another fractionating system using three elements is the Trifid — as the title implies, related to the Bifid. The fractionation/substitution process is effected by means of a table 27 wide by three elements deep. This permits of all permutations of the numbers 1, 2 and 3 (equal to 3 times 3 times 3, of course) in a progressive form over which a keyword alphabet is inscribed. To make up the required 27 elements, a # sign is used — usually appended at the end, but possibly inserted after the keyword.

```
E X T R A O D I N Y B C F G H J K L M P Q S U V W Z #
1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3
1 1 1 2 2 2 3 3 3 1 1 1 2 2 2 3 3 3 1 1 1 2 2 2 3 3 3
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
```

The plain text is written down (without word spacing) and the 3 digit equivalent from the table is written underneath the plain. Then, on some prearranged or indicated cycle, the numbers are taken off horizontally in threes, and the resultant numbers replaced by the assigned letter of the table.

The following uses a period of ten:

```
t r i f i d s a r e | f r a c t i o n a t : e d c i p h e r s
1 1 1 2 1 1 3 1 1 1 | 2 1 1 2 1 1 1 1 1 1 | 1 1 2 1 3 2 1 1 3
1 2 3 2 3 3 2 2 2 1 | 2 2 2 1 1 3 2 3 2 1 : 1 3 1 3 1 2 1 2 2
3 1 2 1 2 1 1 2 1 1 | 1 1 2 3 3 2 3 3 2 3 | 1 1 3 2 2 3 1 1 1

Cipher: E Y M X V U C R Y Y Y Y E A Y V Y O V V X I T D P A T H E
```

You can see the first three numbers horizontally, 111, represent E on the table, and the next three (211) represent Y, etc.

In the example at hand, they give a trifid with the crib "KORE ADECIDEDT HATTHERUN" and (conveniently for the solver) have broken the cipher text into cycles of nine. (The appearance of the worksheet on a cycle such as 7 or 8 seems quite different, and there are slightly different aspects to crib placement, but the reader can easily adjust to whatever cycle is required.)

The first group of cipher was LULSRLXFF, and the next NLT#IWYXR, but I am going to prepare the worksheet on three levels, which on a cycle of 9 means that the letters are under each other. (On 7 or 8, mentioned above, the letters would naturally be offset either to left or right.)

```
L U L N L T N L R # A N L C L
S R L # I W O R V A K K D O R
X F F Y X R V O U K G Q Q G V

B L R N B B C L V W Y B B N W
G L U R # Q C O E X O W Q X W
M F N L W S Z J N P I F X Q T

W N N C L U M U T R N L B G T
I K N V O R L N V U H L K S T
P V Q W X F T W B V U F K I W

      N N R |# N |
      N O O | E W|
      K G Z | E |
```

(If the cycle is other than a multiple of three, it is a good idea to mark off the division by vertical lines, but in the above case it is unnecessary except for the last group of five letters. Remember to save room for three digits to be assigned to each cipher letter.)

Placing a crib can be more difficult than in Bifid, which has only two elements to "accidentally" agree in the case we told you about where an E on the bottom row just before an E on the top row must result in a plain E at the second position of the pair. Remember? In Trifid, we can take some advantage of this property in widths of 4, 7 and 10, which give situations such as:

```
Q..D..E
..B..E.
.D..E..
```

in which there is a two out of three chance that the fourth position in the segment above is a plain D and a certainty that the last position is an E. (Play around with your pencil if you don't see why this is so.) No matter what trinome you assign to E, that will be the plaintext in the last position above.)

In widths of 3,6,9, etc., only in one of 27 cases will the entire plain trigraph be the same as the cipher trigraph —but when you see such a case, you may recognize it. Of course, usually you can ignore this possibility, since in the problem given, the text is unlikely to start with a plain LSX.

In widths of 5, 8, 11 and so on, there is tricky one-to-one correspondence of a one-out-of-nine chance of the plain and cipher being the same letter — but you have to jump around a bit to find "hits." Consider the following diagram where the plain is numbered consecutively and the cipher positions are numbered according to the position of the plain letter which has a 1 out of 9 chance of hitting. The "box" shown is for an 8-wide pattern, but you could figure out the correct pattern for any width wherein the values on rows 1, 2 and 3 of your writeout form a diagonal to the right

```
R  D  A
 Q  P  C
  V  D
```

rather than to the left, as in the DBE and EEE diagonals in the preceding 7-wide illustration. (Numbers below are position, not the numerical value used in encipherment.)

```
Plain:    1 2 3 4 5 6 7 8
Cipher:   1     4     7
            3     6     2
              5     8
```

It is obvious that positions 1, 4 and 7 of the cipher fall directly under the corresponding plain positions, and therefore the first digit of the trinomes must be the same. There is a 1 out of 3 probability that the 2nd digit of the 1st cipher is the same as the first digit of the 2nd plain, and a similar chance that the 3rd digit of the 1st cipher matches the 1st digit of the 3rd plain. Hence the multiplication of probabilities to arrive at the one-out-of-nine chance that the letters of positions one of both plain and cipher will agree. Positions 4 and 7 are easy to understand, also.

But why is the fourth cipher letter (numbered 3 in the above box) supposed to hit with the third letter of plain with a better-than-random chance? Because the SECOND digit of each trinome should be the same. A little pencil-pushing will show that the 5th and 6th cipher letters might hit with the 6th and 2nd plain values for the same reason. Finally, the 7th cipher has the same last value as the 5th plain, and the 8th cipher might hit with the 8th plain for the same reason.

Thus one concentrates on the 1st, 4th, and 7th positions of a crib on a seven-wide box, but can jump around to find one-to-one correspondence on an 8-wide write-out — or an 11-wide.

Sounds complicated, but is actually easy enough if you have grasshopper capabilities. You hope to find a place where more than a random number of plain letters hit at the right places with cipher values.

But in the problem we are about to solve, while we are looking for a spot to fit in the crib (and they were kind enough to strongly hint that the ADECIDEDT and HATTHERUN would be placed in the cycle positions exactly) we notice that in the final positions of the sixth cycle-group of 9 letters the trigraph is RUN — reading vertically, of course. Hoping that this is the right place, we put in the crib:

```
                                        KOREADECIDEDT
L  U  L  N  L  T  N  L  R  #  A  N  L  C  L
                              2.12222 2   2 2
S  R  L  #  I  W  O  R  V  K  A  K  D  O  R
                              2.1112.12.11212
X  F  F  Y  X  R  V  O  U  K  G  Q  Q  G  V
                              ..212 12 ..2 2
```

```
HATTHERUN
B  L  R  N  B  B  C   etc.
222 2212222
G  L  U  #  Q  L  C  W
  22 21 2
M  F  N  L  W  S  Z
 1    2222 2
```

Now we have to arbitrarily assign numerical values of 1, 2 or 3 to the three positions of each plain (and cipher) value. Since the first letter of the crib happens to be associated with a repeated cipher A, and has the advantage of A being a good plain letter, I will start by assigning K the value of 112. Note that I cannot safely call it 111 yet — although the cipher A causes a repeat, we know nothing about any other letter.

Putting in all the information possible, I start to build up two tables, one of the normal alphabet in which I will record the values as they come up, and one of an arbitrary listing of the 27 numerical combinations:

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
2       2           1 2   2 2   1 2   2 1
  2   1 2       1   1 3 1 2 1   2 1   2   2
1 2   2     2       2 2   2 1     2     2
```

```
1 1 1 1 1 1 1 1 1 2 2 2 2 2 2 2 2 3 3 3 3 3 3 3 3 3
1 1 1 2 2 2 3 3 3 1 1 1 2 2 2 3 3 3 1 1 1 2 2 2 3 3 3
1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3 1 2 3
  K                   O R     N     L
```

You will note that L is 2.2, while N has been determined to be 222 and R is 212. Now you know where a 3 has to go. With this to get you started you should have a field day — particularly if you decide to try "running" (and the partial values check.)

In case you didn't think of it, the word in front of KOREA would possibly be SOUTH (more likely than NORTH) and the former is possible while the latter is not, according to the values.

Remember you will have to fiddle with the 1-2-3 arrangement in all probabilities when you start to recover the keyword alphabet, but don't panic. You often will find that an order such as JLK will tell you to switch digits.

If you finally decide on a 312 or 231 or whatever might be the final order, there is little point to changing any numbers —just consider it as an equivalent solution. And one time out of six you will have stumbled onto the correct order by chance.

Tri-Square

The name of this system, though entirely correct, may prove a bit misleading to the novice, since the substitution involves plain digraphs which turn into cipher trigraphs.

Solution can often be a very interesting challenge, but you may end up by putting this problem in your "Favorites" collection.

Any experienced analyst will probably develop his own techniques —and is probably used to the concept that if one knows how a problem was enciphered, one can work backwards and figure out a way of reversing the steps.[1]

Visualize three keyword-mixed squares arranged in some such manner as this (and don't be alarmed that the "book" has a slightly different arrangement (apart from the dropped line in one square.)

```
N S F M U   P A S T I
O A G P W   N O Q R M
V B J Q X   L Y Z U E
E C I R Y   K X W V B
L D K T Z   H G F D C

            R E A D I
            N G B C F
            H K L M O
            P Q S T U
            V W X Y Z
```

The first letter of the plain digraph is found in the left-hand square (which we may call Square I) and the second letter is found in the lower square (which I think of as Square III.) The intersection of these two points in the "middle" square will serve as the middle letter of the final cipher trigraph, but the enciphering clerk has a free choice for any one of five letters as the first cipher value and any one of five for the third cipher value.

This is due to the fact that the clerk can select any letter of the column in which the first plain letter appears, and any letter of the row in which the second plain appears.

Thus, TH plain could be any one of 25 cipher trigraphs, such as MHO or RHL or even THH.

You can see that the plain-cipher transformation of a message might start out:

```
Pt:   th  re  ek  ey  sq  ua  re  su
Ct:   RHL QXR LXO EVZ BAP XSE RXD DIU
```

although only the middle letters H-X-X-V-A-S-X-I would be constant regardless of the whims of the cipher clerk.

This given problem has a crib conveniently placed for you. But while we are at it, we should point out the weakness of the Trisquare consists of the fact that (assuming the enciphering clerk really chooses the letters at random) there is a one-out-of-five chance that the first plain will result in the first cipher value, and similarly, if the last letter of a plain digraph is E there is a one-fifth chance that the cipher will be E.

However, a conscientious cipher clerk may try to avoid this repetition and deliberately choose one of the other letters in the pertinent row or column, and in the given message you will find only one "hit" in the area given, which consisted of 27 letters. (You might expect 5 in such a Tri-square stretch.)[2]

If a crib is not placed for you, there may be a repeated digraph in the plaintext on the even cut, such as TH AT TH ER E.... A little reflection will suggest to you that a repeated digraph must have the middle letter of the resultant trigraph in common, even if the first and third cipher elements did not turn out to be the same. So you would merely try such places as ARD KOX CRF, since the middle R would "fit." Naturally, a longer repeat (on the cut of course) appearing in the crib is duck soup.

In any event, the crib has been placed for you, and the title "Move over, Miss Marple" may at some time during the solution serve as a clue. (Actually, I have added a couple of words to the crib they gave, since the reader who is new to Tri-square might have a little trouble extending enough to make the problem relatively straightforward.)

```
SSS TPC HKO FMK ADD NLP MPN OII LGY KAX XFZ RSD DPH BGW IAI QGV

HUB ADE YCP GPS USC SSN NGD LDE MLS VIZ DKC TQH ESN YQU OIM QKQ

XRD KKE IBS HWT UPR SRG GIM RIY BFA MPC CSO VZU PNV DPN AKD LBE
```

```
SFS ZGW EZC KAP YKP QIQ WMF UGN TKZ NBH VMD OXE AIS DMX LOT ZER
                                s u g g e s t s a n e w h y p o

RXG EMC KTN FTK HUD ILY QZE BHV UKS AOC CGI VGN SVM THY LOD ZEB
t h e s i s w h i c h m u s t b e c h e c k e d o u t

OFZ BXG KUE IRS MMC PSH BXK DIO WMF YON XPD SDP TQG OFE EFS FGA

QSC RSN UCR GGD VIQ ZMI BSS HHA SQH DSC AAN QID OQG COF UGE
```

Incidentally, when I copy out a message such as this on my worksheet, I keep the first and third letters of trigraphs on one line, with the middle level written above, which makes looking for critical middle letters somewhat easier — thus:

```
  K     B     M     X
T Z   N H   V D   O E etc.
s u   g g   e s   t s
```

One of the first things to do is to write down the associated letters of the first and last square, since you have information of what are in the same column of the first, and in the same row of the last. Thus:

```
In same column of No. I          In same row of No. III

S   N   V   L   Z   F            Z  U  T  Y  M
T   G   E   H   P   W            H  G  K  I
R       D   I                    D  S  E  N  C
O       U   A                    X  W
B       Q   K                    R  O
                                 V  B
```

At first glance, it seems possible that the last square is made up with a keyword followed by the rest of the alphabet in normal order, but one must be careful, since such a case sometimes looks very similar to a spiral, because in the latter situation the second row has 4 values in alphabetical order (with perhaps some skips) and the last row has 5 values in reverse order. Even the third and fourth will have three letters that may look interesting in that respect.

An experienced analyst would recognize the Z U T Y M as indicative of a spiral, with the third row probably M Y Z u t or perhaps M Y Z t u. But the best approach at present would be to use a fair amount of paper (rather than confine yourself to three exact squares) and get an idea of the middle square, as well.

It looks rather ungainly, but one could start a long worksheet with the columns of the first box as unassociated clusters, the idea being to bring them into shape by means of association with a middle value. Thus:

```
S E                                   K
T
R
O
B
  V
  D
  U
  Q.....etc.
```

and

```
d  e  s  C  n
o                 R
M  Y  Z  U  T
L  X  W  V  B (Note the logic
K  I  H  G     of placement.)
```

You will see how the middle K ties in the EC digraph, which begins to make a likely-looking third box. "Descendant" is about the only keyword that looks good, and it fits in exactly with what we have. Why not make a complete square on that basis, and start filling in whatever bits and pieces we can in square one, and the middle box?

You can probably finish this yourself, now — with one extra hint for tackling plain digraphs when you have partial squares. Look at the first cipher SsS. Even though we haven't placed S in the middle square, we know that the first plain must be S or T or R or O or B and that the last member must be from the DESCN row. Similarly the next plain must be one of 25 digraphs using TSORB for the first and DESCN for the last "one of five" possibilities. To make things easy for you, I'll suggest the word RESEARCH — and let you go to town.

NOTES ON CHAPTER VII

(1) I seem to harp quite a bit on the fact that would-be cryptanalysts should not be afraid of the dark. I know of several relatively experienced "solvers" who can sail through problem after problem — that they know, and have been shown how to tackle. They know that if they write out the worksheet in such and such a way, and look for the crib placement in the way they have been taught, and put the letters and/or numbers in a certain specific arrangement, they will end up with plaintext.

But suggest that they tackle some problem that they have not been guided through, and the first question will be "how do I go about it?" The fact that a very lucid explanation of the cryptographic procedures of encipherment is given in "the book" doesn't seem to help them in the slightest.

I suppose some individuals don't like *Problem Solving* in the general sense of the title, and shy away from the "most fun" part of all — figuring out that if this is the way the enciphering clerk went from plain to cipher, such and such a procedure would have to be followed to exploit a certain weak point, or provide a wedge into the system, or place a presumed crib, or show up a critical frequency count.

One hopes that more and more of the would-be analysts will work hard at becoming real "solvers" and not just predigested stepfollowers — getting away from their present job descriptions as cipher clerks rather than cryptologists.

If you hear the snap of Simon Legree's whip, forgive me for my seeming intolerance, but I live in hopes that certain lazy thinkers will be shaken into giving Poirot-designated cells a little workout.

(2) Such suppression of hits can boomerang, since anything that throws a monkey-wrench into the expected probability of occurrence could offer an entry wedge. Don't forget the weakness of the ENIGMA machine used by the Germans in WW-II, which dictated that an E plain could never be an E, etc. Placing a crib was occasionally possible because of the properties of the gadget, but a similar situation could conceivably occur if you suspected a conscientious clerk was deliberately suppressing certain relationships.

Chapter VIII

OTHER CIPHERS

A CONVENIENT source for discussion of various ciphers is the output of the A.C.A., which has developed a lot of fun-problems that may be unique to their membership. While they have their own names, such as Tri-square, Grandpre, Trifid, Beaufort, Ragbaby and so on, the principles are not necessarily new, and usually are derived from some twist on standard substitution or transposition procedures. Therefore, they present good material for diagnostic study.

While some are moderately difficult, there are naturally no systems which could satisfy the requirements of the most under-developed government, since the problems are basically geared for paper-and-pencil solution by enthusiastic hobbyists, ranging in skill and aptitude from rank beginners to a few ex-professionals.

It would be a simple matter to make some of the problems very difficult, if not impossible to solve by the average reader. As I have underlined many times herein, a Playfair or even a simple substitution without word-length can be quite a task when there is no crib given.

But that would defeat the point of this whole enterprise, which is not to invent the unsolvable system but to tease and beguile the willing victim, giving that person the opportunity to figure out "how to pick it up" — and perhaps have a little fun with an apt solution.

No machine systems as such are represented,[1] but as the reader may know, almost all cipher machines — at least the simpler commercial models — can be represented by a paper-and-pencil system.

A wheel may turn out to be a square or combinations of squares (on paper) and a fancy series of stepping switches may be thought of as merely (?) a way of determining which (known or unknown) alphabet applies at a given position.

Baconian

Some of the systems we will discuss are classic examples, such as the Baconian cipher.[2] This is not really an encipherment so much as it is a game of hide-and-seek. Letters consist of five binary elements — A and B being the conventional means of distinguishing between them.

The letter A is always AAAAA, B is AAAAB, C is AAABA, and so on, by logical binary progression, ending with Z being BABBB. I and J are equivalent (ABAAA) and the U/V combination is BAABB.

You can dream up your own way of separating the universe into two categories — vowels and consonants, odd and even, black and white, or any combination that lends itself either to literal text or to a pictorial representation.

With vowel/consonants, you could swing either way — with ABOVE being L or DEVIL being L, depending on whether you choose vowels to represent A or B. (Naturally you wouldn't want to switch horses in the middle of the Potomac.)

A brick wall may be shown in a drawing, wherein some bricks are shaded and some are not. There you have a valid distinction, and a method of disguise.

Since there is no such thing as solving a monoalphabet, or discovering the cycle of a polyalphabet, the "fun" (and not all solvers like this sort of amusement) is in search-and-find missions. And they could be made ridiculously impossible, or relatively simple.

Suppose I say that I will send a flock of random (?) 5-digit groups, bearing a message: 18070 02711 82464 83162 29085 16685 27291 90875 68636 44973.

Short samples might be tough, but yet if you looked at the frequency of the digits I used, you would probably get the answer. Remember, there are 16 letters that begin their patterns with A, and only a few of the remaining B-beginners are common. More importantly, the rhythm of "elements," whatever they may be, is most important.

I make no apology for omitting a long sample for us to jointly solve.

Gronsfeld

The classic Gronsfeld uses a numerical key on literal text. The easiest way to think of it might be as an additive system, in which the plain is moved along a normal sequence by the amount indicated by the keying digit.

Thus (depending on your convention) if the key is 4, a plain of T would be cipher X, and a key of 9 would make A become J. (Some conventions allow for only 9 values, with 1 holding A to A, and 2 making it a B.)

The weakness of the system is apparent, in that no plain value can be enciphered by a letter removed further than nine letters higher on a normal alphabet. Therefore, E is always E, F, G, H, I, J, K, L or M (and N, if 0 to 9 is added). Thus THE would have a maximum of 1000 ways of encipherment, instead of 17,576 — which may not seem like a big deal, but is actually very important.

You can easily see that if a crib of THECATISON is given, you could place it by merely searching for an area where you found a letter from the "possible" set of TUVWYXZABC (depending on your rule for adding 0, if any) followed by one of HIJKLMNOPQ and then one of EFGHIJKLMN, and so on.

I'm sure I need not dwell on methods of placing cribs, reading depths, and so on. To make matters even more simple, some versions consist of a periodic cipher, with a digital key repeated over and over again.

It is theoretically possible to read a message using entirely random key, where there is no "depth" or manner of key construction to assist you, but there are likely to be manifold ambiguities, and I would not recommend this for a quiet afternoon's occupation.

To make sure you understand the mechanics, place the crib PHENOMENA in the following. (You can easily read the text, and after you do so, you may discover something odd about the key.)[3]

GKQHN QLWQM VQSLT AUDBH MBHNK RUSLQ MRWOW BOJXV TYCKT P

Porta

Another system with an inherent weakness is Porta, also a classic method of encipherment (by which I mean you will find it in all the literature, as having been used for centuries.)

It uses reciprocal alphabets in which the low values (A to M) can be slid against the high values (N to Z) in only 13 ways. Thus if B is P, P is B, and C is Q — and P can never be a value out of the A to M range. ACA uses this as a periodic cipher (repeating a fairly short key, as in the Gronsfeld above). In actual usage, one could use any literal key — from a book, a specific page of the New York Times, or some other source.

The Key "name" is arrived at by writing down a tableau of the 13 possible slides of N to Z underneath the A to M sequence, and then calling the first one A/B, the next one C/D, etc., with the last being Y/Z. Therefore a key of ADVISOR uses the first juxtaposition for the first substitution, the row being OPQ... for the next, the row beginning XYZN... and so on.

However, even though the key might be carefully used, and unpredictable, you'd be surprised how easily you can place a crib, or guess at phrases when the general tone of the conversation is known. Of course, I may be understating the level of difficulty when the key is truly random and not repeated at any time. (As you might guess, "depths" are the bread-and-butter of cryptanalysts and the bugaboo of cryptographers.)

You will appreciate the fact that Gronsfeld, above, had only 10 ways of disguising any given plain letter, but Porta has 13, which seemingly small distinction makes a big difference when it comes to ambiguous "solutions."

Just to test you, could you read the following if I said it was the beginning of a Nursery Rhyme? (To make it easy, I used the keying procedure in a very elementary way.)

W N P X N A Q W V Y Z I S M F G B F V S W X N N E
M U T E R X Q M Q Y O L V G Q C V A etc.

If there is an occurrence of THE, it must be a low-high-high trio. I see several such, but at one of them you will get a clue as to what might be happening. (I'll say AND appears, as well.)

Gromark

Gromark is a system which we will not go into deeply, but which uses a method of keying worthy of mention. In the Gromark, encipherment is effected by means of a Normal plain set against a Mixed alphabet. However, instead of cycles or predictable slides of the cipher component, one finds the plain value on the top (normal) component and counts along a specific number of positions to the right, then takes that letter in the cipher alphabet, immediately below.

The choice of how far to count along the sequence is determined by a digital key. In other words, one is "adding" 0 to 9 to the plain value, just as in Gronsfeld, but it is on a mixed sequence, set underneath a plain sequence.

If it were done by "adding" on the mixed sequence ONLY, without reference to a normal plain sequence, it would be much tougher. The reason for this is that, in usual manner described above,even though plain text A might go (for example) to one of the ten letters on the mixed sequence KTDINEWXAJ, the plain B must go to one of TDINEWXAJB, plain C going to DINE etc., with relationships easy to establish.

But if you stuck to the one sequence KTDINEWXAJ etc., there would not be any necessary relationship between what plain T goes to and what plain U goes to; since the reference points of the normal would not permit certain important relationships to be built up, you would need a long, long crib, or some depth which might permit of solution.

It would be easy to make the problem harder by using a short cycle, as well, perhaps as a numerical constant (such as a second step of adding 9 8 3 6 5 over and over to whatever digital key you start with.) But we promised to give you things that are fairly readily analyzed and solved, rather than tell you how to make things miserable for the cryptanalyst.

Back to the recognized Gromark. The components never change their position relative to each other — therefore every plain value has 10 possible cipher representations, as in Gronsfeld — as we said.

What makes it more difficult, however, is the fact that the cipher sequence is mixed (rather than normal, as is customary with Gronsfeld) and the numerical key is derived rather than repetitive in some cycle.

The derivation comes from a process with which you may be familiar, if you have taken the right math courses — a Fibonacci series. On some cycle (frequently five-wide) the key is derived from a starting group, by adding the first position to the second and placing the result in the sixth position.

Similarly, positions 2 and 3 are added to make position number 7, 3 and 4 to make 8, and so on. (All addition is non-carrying — a practice in cryptography so common that I sometimes forget to point it out.

For example, if the "starter" group were 48671, the key would run as follows:

```
48671 24383 67119 382021 etc.
```

We will not dwell long on methods of solution, which are fairly straightforward. But it might be worthwhile showing how to place a crib. Suppose you have a stretch of cipher and the accompanying key which includes:

```
7 7 2 6 6 4 9 8 2 0 3 7 0 2 3 0 7 2 5 3 7 9 7
J C N W Z Y C A C J N A Y N L Q P W W S T W P
```

and (not knowing the cipher sequence) you are told that the word SUBSTITUTES appears — and for this example, in the stretch which runs from the beginning J to the final P above.

Since the Plain sequence is normal, a repeated cipher letter, with different key numbers above it, must stand for plain values removed from each other exactly by the difference between the two numbers. Thus the stretch C A C with keys 9 8 2 above it implies that if the first cipher C is M, for example, the second C is seven positions to the right of it on the plain sequence, or T.

```
J K L M N O P Q R S T U V W X
              C
```

If the first C had been O, the second would be V, and so on. In other words, if a letter is placed on the cipher sequence, one can decipher all ten possible key indications for the cipher C, yielding ten consecutive plain values — M to V, for example, some of which (such as Q) might not appear.

Now there are two ways to skin this cat named Solution, given a crib, cipher and key. If the cipher repeats at a reasonably close interval, such as in C A C in the stretch of cipher above, you can easily determine that the plain must be 7 apart — M-T, N-U, or O-V, etc.

So all we have to do is ask ourselves, does the crib contain a difference of 7 two letters apart?

In order to check such a thing quickly, it is wise to prepare a little table of differences, on a sort of triangle, in which the first row represents adjacent letters, the second row two apart, and so on. (I have written the the numerical value of the plain letters for your convenience, although all solvers will soon remember that E is 5, J is 10, O is 15 and T is 20, etc. These 5-10-15-20 anchor points help, if you are trying to remember whether Q is 17 or 18.)

```
                19 21  2 19 20  9 20 21 20  5 19
                 S  U  B  S  T  I  T  U  T  E  S
Adjacent diff:    2  7 17  1 15 11  1 25 11 14
   x-x              9 24 18 16  0 12  0 10
   x--x               0 25 7 etc.
```

Now all we have to do is find a favorable case where the differences in cipher repeats matches the plain differences, at the correct interval. (Note that to match differences, we must measure them in one direction for the plain, and the other direction for the cipher.)

Since I have arranged the table by subtracting the left hand letter from the right, I must look at the cipher in terms of subtracting the other way — the right from the left. If this confuses you, work it out on paper.

Incidentally, in addition to remembering the values for the letters A, B, etc., you had better get used to remembering differences modulo 26, and complementary differences. If A to R is a difference of 17 (1 from 18) then R to A must be 9 (17 + 9 = 26).

Adjacent cipher matches are the easiest thing to look for, and give the widest range of the crib matching (since the triangle gets smaller as letters get further apart). However, the 23 which the two W's give does not appear on the adjacent difference list. How about x-x?

There is a difference of 7 with the C-C hit, but it doesn't appear on the second row of the triangle.

This means that the word SUBSTITUTES must cover the exact area between the A (between C's) and the first W. (In most problems given, you will find concrete evidence of the correct setting, rather than resort to negative information.)

```
7 7 2 6 6 4 9 8 2 0 3 7 0 2 3 0 7 2 5 3 7 9 7
J C N W Z Y C A C J N A Y N L Q P W W S T W P
          S U B S T I T U T E S
```

This is an especially fortunate placement, but you wouldn't need such a benison, since you will note the the N--N would have given a matching difference in cipher and plain of 1. Also, the A---A in cipher (with a different of 1) matches the S---T plain.

To build up the cipher component, write a normal alphabet, and then count along it from any given plain the number of steps given by the key, then write the cipher value below.

Thus, the first value would be obtained by finding S on the top strip, counting 8 spaces to the right, dropping down to the cipher component and writing A.

Similarly, cipher C is is two spaces to the right of the position held by plain U, and so on.

```
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
A J           Y     P               Q W N C L
```

Now you can decipher any of letters A, J, Y etc., by counting backwards the number of steps given by the key, and reading the plain value. Thus, the cipher C ahead of the crib translates to N.

Unfortunately, the cipher sequence seems mixed (in one of the ways to be described later) but you could read enough of the text to fill in most of the values.

You might have deduced that if you don't have a crib, you may still be able to break into the system statistically. Since the numbers associated with any given cipher letter represent a stretch of ten consecutive values along a normal alphabet (such as C to L or X to G) you can prepare a table, with A to Z as the rows and 9 to 0 as the columns.

The reason for the "backwards" numerical sequence is that after you have made checks and tallies for all the entries — D 7, I 3, Z 0 or whatever — you will be looking at a bar graph of a segment of the normal alphabetic frequencies, in order.

It is sometimes easy to see matching frequencies on the diagonal of your chart with letters not in the keyword. You may be able to prove that G H J K L fit one after another at an interval of 1.

You can then combine their frequencies, and spot the fact that this is probably the PQRST area of the normal, or something equally productive.

Again, if you don't immediately see what I am talking about, think of the causes and effects, and work out this simple procedure wherein you may be able to derive a good deal of the cipher sequence without guessing a word of text.

Vigenere, Beaufort and Variant

These names are really connected with the table (or series of alphabets) used to get Cipher (C) by applying Key (K) to Plain (P).

The classic Vigenere goes way back — sort of like a Caesar slide, but using different Key letters, or levels of a 26 x 26 square, at certain times. If you prepared a square, with the normal alphabet on top as a plain sequence and the 26 slides of the normal beneath, you could use the first alphabet (beginning with A) whenever the Key A was called for, the second alphabet (beginning with B) for Key B, and so on:

```
      Plain: A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
KEY A        A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
    B        B C D E F G H I J K L M N O P Q R S T U V W X Y Z A
    C        C D E F G H I J K L M N O P Q R S T U V W X Y Z A B
                    (Cipher)
```

Of course, if you are mentally agile, you can dispense with the square and perform the transfer in you head, since C equals K plus P (where A is zero, B is one, etc.). Thus a plain of D plus a key of B equals E.[4]

The name Vigenere[5] if you see it titling a cipher problem may indicate the method of encipherment is being used with a cyclic key — perhaps as a polyalphabetic cipher, up to 15 columns wide, with the key for each column part of a keyword of that width.

Beaufort is a similar polyalphabetic, but has a different combination rule in which C equals K MINUS P. This has the effect of reversing each row on the tableau (with the Key name still being the letter under plain A.)

```
Plain:  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
Key Z:  Z Y X W V U T S R Q P O N M L K J I H G F E D C B A
  &  Y:  Y X W V U T S R Q P O N M L K J I H G F E D C B A Z
Cip X:  X W etc.
```

Variant is almost the same, except that C = P - K. Since the above are really only variations on the method of combining plain and key and to get cipher, we will not dwell on "problems to solve" (even though an unwary analyst might have difficulty with reading or crib placement until the correct formula was tried.)

Autokey

Lots of people talk about Autokey, but as a general rule no one is doing anything about it. There are applications for this technique in certain communication systems and fool-proof machines, but as a method of preparing a message to go out over the air, it ain't so hot.

One reason for this disrepute is that a garble, a glitch, or just a teeny weeny human error may really foul things up. And there are inherent weaknesses connected with some types of usage.

It sounds like such a great idea. You write out a literal message, and put down some prearranged key for a given number of positions — say MARCH.

Using any method of substitution you have chosen — Vigenere, a mixed keyword strip sliding against itself, or whatever — you encipher the first plain letter with the key of M, the second with A, and so on up to the fifth with H.

What do you do with the 6th cipher? You use the first cipher letter (obtained at the first encipherment) as key. Then the second cipher is used as key for the 7th plain, the 3rd for the 8th, and so on — you never run out of key, since it is several steps back from wherever you are in the plain stream.

So that is cipher-text auto key? Where is the weakness? Well, for one thing, a missed group, a garble, or any communication problem may be quite an aggravation. If you are lucky you can pick up the thread again, since the cipher (actually key) is right there for you to look at.

But someone you may not want to read your mail can look, too. If the sequence is known — as it would be in Vigenere — that person may have picked up the thread long ago, suspecting Autokey.

(Of course, if you use a mixed sequence unknown to the Peeping Tom, he would have a much more difficult problem.)

Since the above was an example of Ciphertext Autokey, what about Plaintext Autokey? The same general procedure for "setting up" prevails, but for the sixth letter (if the primer key is five wide) would use the first PLAIN letter as key, the second plain for the seventh key, etc.

Now if a cipher group is dropped, or a communications glitch occurs, you may really be in the soup. For if you have lost a few elements of plain, you may have difficulty in getting restarted. With a modicum of luck — and skill — you may be able to guess the missing plain stretch, but that puts you in the same position as the eavesdropper.

So let's not dwell upon Autokey in this explanation. In some ways it is too easy, if you have a crib, or it is Ciphertext Autokey, and in some ways it is too difficult for "fun" solving (unknown alphabets, for example, if nothing is known about the plain text.)

There are not many general sources for "cryptograms-for-solving" and since many puzzles appearing in a "club magazine" are unique to its pages, there is no point in dwelling on them all.

Concerning the more advanced systems, including a number of commercial machines, this general overview is no place for a chapter on "how to solve ENIGMA" or "how to recover the pin patterns of a Hagelin model such-and-such."

Remember, however, that most such gadgets still present a polyalphabetic substitution to solve — albeit a very much more difficult problem than is contained herein.

The big difference is the length of the key, which may vary from "long", to "millions", to what some might claim to be "infinity". And if there is no repeated usage of key, and rigid communications security is maintained, even theoretically "solvable" systems can be placed well out of the range of practical attack.

Code-breaking is also not within the purview of such a "beginner's handbook". After all, this chapter expressly mentioned "elementary".

Book Cipher

There is a substitution cipher which looks as formidable as a code, in that fairly sizable groupings may represent a single letter.

By using some "book" (hence the appellation) or other printed material in which a given letter can be easily identified by some such designation as page, line, and word order (and even position within word) you can make a fairly secure (but cumbersome) cipher.

You might confine your choice to the first ten letters on any line, giving the recipient merely page, line and position within that line. Or you could use the entire line, by having a larger number for position designator.

Hence, a letter might be represented by as small a cluster as 3-8-5 or as large a group as 107-31-19.

A classic example would be Mr. Friedman's solution during WW-I of some Hindu "Indian Independence" messages. (See *The Codebreakers* for details.)

Much more recently, we have seen the system exploited commercially by the appearance of certain widely publicized "contests" wherein applicants (for an admission fee) are given enciphered messages purporting to reveal the location of certain hidden "treasure". They are also given a long list of authors, from which they can hope to discover the one book which will serve as the key.

It is indeed a formidable task to break a single message using an unknown book. But if the choice of "key" for the substitution is limited to a single page of text, or many messages are available, or a crib is handy (that's a lot of or's) something can be done — as Friedman demonstrated.

If each letter has to be replaced by a designation for page, line, and word (as a minimum, if you use the first — or better still, the second — letter of a word for the magic replacement) you need at least a 5-digit number, with a hyphen between page and line, and line and position, which makes for a somewhat inefficient ratio between plain value and length of cipher equivalent.

Every Christmas my "card" consists of some variation on a relatively easy-to-solve cryptogram. Recently I have used a form of PseudoBook-Cipher in which I write a greeting, then write an entirely different "letter message" which I use for key to encipher the greeting message. I merely give a string of consecutive numbers (from 1 to whatever) for my greeting, and the solver has to replace "1" with whatever value he assigns to that number when it appears (probably several times) in the easily-guessed key message.

The key sheet has plenty of entry points, since it is divided into word lengths, and oftentimes plain bits of text are interleaved in the enciphered "key" text.

Thus, looking at the key, you see such give-aways as "We visited 8-14-95-2-77 in 3-19-101 66-13-2-112" which quickly determines that the visited party goes —R-, since you can guess we went to NEW YORK.

Obviously, you should fill in the third position of the 1-2-3-4-5... greeting message with N, the 19th with E, and so on.

A close relative to the "book cipher" is a so called Dictionary Code which replaces a word with the page-and-line (or word number) equivalent. Thus, if I were on good terms with my rich aunt, I might send a telegram "629-11 120-6" which would be translated as SEND MONEY if you knew I was using my Merriam-Webster pocket book dictionary.

One can get around the "hyphen" problem by using a letter-for-number substitution with provision for separator — say A to J for the numbers 1 to 0, and K for a separator. Thus my request for money would have been GBIKAA ABJKG, with spacing between

groups. (Or you could also use the separator between groups and send text in sets of five cipher letters.)

To make things a little tougher on the kibitzer, you could use a more complex substitution — at least 11 letters (based on a keyword) or even all 26 letters with several variants for each number.

This would really be a code problem, and we stated that such was beyond the scope of this "elementary" discussion.

Seriated Playfair

A quick word about a variation on the Playfair, which is certainly a "miscellaneous" inclusion, but happens to have some challenge, and is a good example of "how to tackle it."

The ordinary Playfair, a straightforward digraphic system, has a gimmick added — the plain text is written out on some width, in sets of two width-lengths. Then each two lines is enciphered by taking VERTICAL pairs, putting them through the square in the usual fashion. (Check on the first page of Chapter Six, if you have doubts about the Playfair square, or any misgivings.)

Thus, if the text started SERIATED PLAYFAIR IS A SLIGHT TWIST with a width of 6, for example, the plain text would be prepared:

```
SERIAT  YFAIRI  RTWIST
EDPLXA  SASLIG    etc.
```

You will note that the bottom of the first set reads "EDPLxA" — the intrusive X being added to take care of the impossibility of enciphering AA. The rule is simple — if a plain value on the bottom line is the same as the value above it, throw in an X and go on your merry way.

Now the pairs are taken off, enciphering SE as LW, ED as WC, RP as AO, IL as SE, AX as WA, and TA as PK, etc.

The cipher text to be sent would read LWASW PWCOE, etc., the cycle of 6 being disguised by the usual 5-letter transmission groups.

Once you place a crib (or make a brilliant supposition as to text) the procedure of recovering the square is identical to that explained in an earlier chapter, except that one must be careful to treat the text as vertical pairs. Since the pairs consist of plain values removed by several positions in the running plain, some of the statistical approaches may be made a trifle tougher.

But the main problem is the somewhat tedious task of finding the cycle. It is straightforward, since one merely throws out any length in which a cipher value on the top row (of two lines) falls over the identical value. And of course you don't have

to write the cipher over and over, assuming a width of four, then five, then six, etc. Best illustrated by taking a sample problem.

```
TSCNF FHGAS TXZBA PYPLE TRNNI YXIYQ PGUXO RZOZO UZYLX PZCPF
WGMLE GXANF GTNRZ XZOGL SEFNE OXECR XXYEX SEVXT PALZC TPMYG
EHPIE MADAG LXNXM AEUYP ZPNLY RQGLG NKOIL SPXYL
```

You will readily appreciate the fact that the cycle could not be ten, since TSC.. would fall over TXZ.. Similarly, if the cycle were eight, the S's would clash. A simple computer program does this calulation for you, or you can find it by simple industry.[6] In the problem given above, they gave a strong hint by telling you the plain "NTTOTHESAMEEXT" appears. The fact that these 14 letters obviously start in the middle of some word is supposed to get you started on the logical width of seven. (The originator could be being nasty, but this is frowned upon by the editors.)

Of course, in trying any likely spot you may take advantage of the crib, in that there can be no "crashes" between plain and cipher, as well as no crashes from line to line.

In any case, the correct cycle of seven is quickly determined and the text prepared:

```
TSCNFFH  APYPLET  YQPGUXO  YLXPZCP  XANFGTN  SEFNEOX
GASTXZB  RNNIYXI  RZOZOUZ  FWGMLEG  RZXZOGL  ECRXXYE

XSEVXTP  YGEHPIE  NXMAEUY QGLGNKO
ALZCTPM  MADAGLX  PZPNLYR ILSPXYL
```

How to find the crib? (Here is where I argue that budding analysts should think about it, and answer the question themselves.) So we will assume that you said "Of course. All you have to do is look for a repeated vertical digraph in the crib, if you are lucky — or at least a reversal. Or I'll think of something else if I have to."

The crib written out seven wide will fit some place over and below the cipher:

```
NTTOTHE  ent?
 cipher
 cipher
SAMEEXT
```

We are fortunate in having the crib probably given to us at the right cut, but many times the solver will have to try the crib starting at each alternate position—in the case of seven wide, fourteen different starting positions. Of course, many of the possibilities have the same pairing of plain letters—the N over S and T over A would appear even if we moved the crib by a few positions.

And we do have a reversal — the TE in the fith position and the ET in the last position of our two-level crib writeout. Since the originator was kind enough to hint at the width, one assumes that he wouldn't cheat in other ways, and that we must search our cipher writeout for cases of such a reversal. (You will remember that if TH is OD then HT is DO.)

Sure enough, the sixth set of paired lines has an EX and XE where we want such a phenomenon. Since the recovery of text is the same as a usual Playfair, I'll leave the fun to you. You will note that you were fortunate enough also to find two cases of letters being on the same row or column of the square, since NSE and TEX both are offered to you, and must appear as one of the two following arrangements:

```
  T          N
NSE    or   S
  X          TEX
```

Be careful when you try to add the plain TA going to cipher EC. At first glance, one may be tempted to throw out the first possibility because one cannot make the proper "cross" of a four-way entry. But remember that the entries could (like the ones you started with) be on the same row or column, as in

```
  T                N
NSE                S
  X     or         TEXAC
  A
  C
```

The OE going to NX helps materially in getting you on the right track. But it ain't easy, as I realized when I noted that this problem was No. E-26 in the CRYPTOGRAM — therefore probably the hardest, except for some in the Analyst's Corner.

One of the reasons this particular example may give you fits is that (by accident?) a number of the pairs given in the crib fall in the category of same row/same column. Why do I leave in as "an example" such a nasty challenge? Because it will make you very aware of the booby-traps that occasionally lie in wait — and I have made up for its nastiness by giving you many extra hints.

When I tackled this myself, only when I took a flier on the TEXAC above — which could just as easily be EXACT to start a keyword in the square — did things settle into place. Just to make things easier, I'll give you a CAESAR slide of the beginning — HYHUBWL.

NOTES ON CHAPTER VIII

(1) Some years ago, my good friend, the late Eugene "Danny" Waltz fixed up an "unknown cipher" for a special article in *The Cryptogram* which just HAD to be based on a commercial machine known as the KRYHA. (Danny, as he was known when he worked at the same "Puzzle Palace" that nurtured me, got his nickname because of his "nom" of DANCER. (His first name, though few called him that, was really Eugene.) Waltz— get it? It became DAN SURR in later listing.)

Anyway, I couldn't resist the inclination to write an answering article explaining the source, since there is an interesting pattern of "jumps" between settings on what is a fairly complicated polyalphabetic cipher using slides of a mixed alphabet against another alphabet.

And I have seen an artcle by O'PSHAW giving a rather thorough analysis of KRYHA — so perhaps we are entering a new phase in the level of the amateur cryppie's approach to "worldly" problems.

It may give me the opportunity of recounting one of my first experiences in the pre-War days, when I figured out a way of short-circuiting the cumbersome practice of taking seventeen different cipher counts of a long KRYHA message, looking for the one which was monoalphabetic.

(The trick was to write the text out once, in a specific format, then slide a specially constructed grille through 17 settings, one of which should "look" monoalphabetical enough for serious consideration.)

Since I was a relative tyro at the time, never having been even exposed to a cipher machine, I use the story as a word of encouragement to those solvers who are scared by the awful words "machine system."

As I said, many machines can be turned into the equivalent of a hand system — albeit using cumbersome (and voluminous) work sheets.

A further "lesson" relating to the early story. After licking the machine by one method or another when straight plain text was enciphered, we found ourselves in the enviable situation of reading ENCODED messages because of our knowledge of alphabets, indicators, etc.

Tackling the encoded messages from scratch would have been wellnigh impossible, even though we had reconstructed the code. This emphasizes the importance of certain rules of cryptographic security. You can supply your own "DON'Ts."

(2) There is an interesting connection between the Baconian cipher and the Friedmans — both William and Elizebeth. She was hired by Col. Fabyan (of whom you have read some juicy tidbits herein) to work on the project headed by Elizabeth Wells Gallup to "prove" that Francis Bacon was the real author of the Shakespeare plays. William was basically hired as a geneticist, but took a specific interest in certain areas of the Baconian investigation.

Not coincidentally, the future Friedman pair naturally saw a good deal of each other, and one suspects they each were attracted by and appreciative of the obvious talents of the complementary member of the soon-to-be team.)

It did not take the couple any great length of time to realize that the fabric being woven by the sincere but very misguided original investigators was gossamer. And when WFF was able to take their very methods of "proof" to create almost any fabricated text one wished (including Friedman-proclaimed authorship of the plays) the pair could not in all conscience lend their serious attention to such a project.

A fascinating tale of the demolition of the "Baconian" theories is told in *The Shakespearean Ciphers Examined*, written jointly by the pair and published in 1957 by the Cambridge University Press.

I treasure the inscribed copy that the Friedmans gave me when they visited me in England in 1957, while on official business there. (For a couple of years I was the Deputy to the Senior United States Liaison Officer in London, and in charge of the American contingent working with the British Foreign Office at the Government Communications Headquarters in Cheltenham.)

(3) You undoubtedly noticed that the key was "self-generating" in that the stretch 798826760 you should have obtained at the crib placement has a property in which the first and second numbers add to the sixth, the second and third add to the seventh, and so on. This Fibonacci series is used in the so-called Gromark system, which you will meet with herein.

(4) There is no universal agreement as to whether you should call the letter A "0" or "1." I "grew up" in the Army shop, which calls A "one" (usually). The Navy prefers "zero." This confusion may be due to the fact that in certain cipher machines, the "arithmetic" is simpler with the A/0 assumption.

I find it easier in dealing with Vigenere, as in the example shown, to use the C = K + P formula by adding the values on the A = 1 basis, and then subtracting 1. Thus Plain D (4) plus Key B (2) equals 6-1, which is 5 (E).

(5) I have been waiting for an opportunity of referring the reader to one of the most comprehensive books in the field of general cryptography which I know: *The Code Breakers*, by David Kahn, who has become one of the foremost authorities in the field, despite the fact that he is not a card-carrying member of the official fraternity.

Although early in his life he fell in love with the subject, and was in fact assisted in his search for information thereon by no less a master than William Friedman, his life was rather as a newspaperman, outside the portals of such places as NSA.

Kahn is a very active writer, an excellent speaker, and the living exception to my first stated principle that most writing in this field is by someone who necessarily can't know all that much about it. He is also a very personable gentleman, although his enthusiasm and dogged sense of "following the story" may have offended some of the Rajahs on the inside of the business.

The Codebreakers — all 1164 pages of it — is crammed with information ranging from "The first 3000 years" to "Messages from outer space."

I can not point to any sentence he has written on matters of which I had some inside knowledge, since my Agency had an official policy of "no comment", and saying "This is right" or "This is wrong" is tantamount to publishing a true statement, whether it be the table of organization of NSA or a much more secret matter.

Actually, I saw the book in its "proof" form, since Kahn bent over backwards in his attempts to get official blessing, and gave a copy of the manuscript to the establishment for "vetting" (as the British say).

Regardless of whether or not I think he guessed right about certain phases of our war-time work, this book has more to offer in the way of who-was-who, what-is-what (and even a little of who did what to what) than most others ever written on the subject.

And if you want to read all about Mssrs. Vigenere, Trethemius, Porta and company — or for that matter our own Yardley and Friedman — I can recommend no better source.

A slightly juicy tidbit that I can slide into this slot, relating to *The Codebreakers*.

I have been fortunate enough to escape having my name appear in most of the exposés about our heinous activities. Several near misses, such as the fact that James Bamford's *The Puzzle Palace* wasted some paragraphs on the questionably exciting news that John Larkin left Southampton in 1956 after serving a term as D/SUSLO in Cheltenham, England. Guess who replaced Larkin?

However, in the case of Kahn's book, I did barely make it — but only in a footnote 'way back on page 1091. One of the first security "trials" took place in 1954, when Joseph Peterson was "made an example of" after he passed certain information on to the Dutch, our former ally.

Joe was a very misguided, idealistic type — who during the war worked closed with Col. Verkuyl of the Royal Netherlands Indian Army. When "peace was declared," quite logically things changed a bit — and the Dutch were no longer necessarily privy to the type of information Peterson might have been in the habit of (quite "legally") chatting to Verkuyl about.

Peterson, a very religious (but somewhat odd) personality, seemed to put himself in the God category, and decided that what was good for our country in 1945 would necessarily be worth preserving after the war.

Since I had known Pete (as most of us called him) for years, and he had quite capably served on my Technical Staff for a long period during the conflict, I appeared before Judge Bryan to answer a few questions.

I could not help but reply to a query concerning the possible damage to the cause by the (to us) very new and strange publicity. I lamented the fact that until "now" even my teen-age sons had no inkling of what I did for a living (other than that I was in some sort of hush-hush activity) but due to the newspaper accounts, now probably all his school chums were kidding him about spies in the house.

Well, this made the day of certain reporters, and the headlines of our evening tabloid read "NSAers never tell their wives."

Also, my penchant for getting close to the wrong people got me in a spot of trouble with my then-boss Kullback shortly after the Martin and Mitchell case broke. (I am reminded of age when I realize that many readers might not recognize the names of the two cryptanalysts who defected from NSA.) Guess who Martin was working very closely with, when he skedaddled to Russia, by way of Mexico?

I'm sure Kully wouldn't have faulted me for using the splendid programming talents of Martin (he did an outstanding job) but to make matters pretty awful, I had (just before the balloon went up) pursuaded Kullback to write a nice letter of commendation for the soon-to-be-culprit.

My boss accepted the buck very stolidly — and never passed on one word of annoyance at the embarrassment I had caused him.

To this day, I would have to admit that if I didn't know any of the facts which obviously came out later, I would have no hesitancy about letting it be known that here was an outstanding brain, and a fine programmer who had done excellent work for me — and I would hand him the next assignment.

(6) A small time-saver that helps in many searches: Quite often you can test two or three widths at once if you don't mind the eyestrain of following a diagonal. In the problem given, if you wrote it out 5-wide you could check the 4-wide possibility by comparing letters on a diagonal to the left, and a 6-wide possibility by looking on the slant to the right.

Thus:

TSCNF	RZOZO
FHGAS	UZYLX
TXZBA	PZCPF
PYPLE	WGMLE
TRNNI	GXANF
TXIYQ	GTNRZ
PGUXO	

In this particular case (Seriated Playfair) you have to stick largely to comparisons within sets of two — 1st and 2nd line, then 3rd and 4th line, etc.

You can easily see the "crashes" in the same column, such as the TT in rows five and six, which knock out the possibility of a width of five. But rather than rewrite the blocks, you may be able to see some hits on the diagonals which would knock out a width of four or six.

This technique is of less value in Seriated Playfair since you have to be take the "two rows at a time" into consideration, which means you have to be aware of the exact numerical position at all times. For instance, the A on the second row and the A on the third row, even though on the diagonal to the right, are not a conflict here, since the A's appear in different clusters of 12, if one is trying to knock out a width of six.

However, the positions of N in the GXANF GTNRZ groups (59 and 63 of the total count) fall within the set of eight letters which knocks out a width of four. You can probably spot such conflicts just as easily by running your eye along the text, and considering the intervals between "hits," making a quick check on whether they fall within (good information) or across (no valid info) the span of any width being considered.

But if you were looking for significant repeats on a Quagmire, or perhaps sizing up the monoalphabeticity of a column on a certain width, looking (or counting) down the diagonal is helpful, and saves more than half of the work of re-writing. Perhaps not a very exciting "hint," but a small technique of procedure.

Frank W. Lewis received an Exceptional Civilian ribbon in 1969 from Admiral Noel Gaylor, who was then Director of the National Security Agency.

Lt. Gen. Marshall Carter, DIRNSA, in 1966 presented an award to Frank Lewis for exceptional service. Dr. Louis Tordella, D/DIRNSA, is in the center.

Lt. Gen. Gordon Blake, DIRNSA, is shown in 1964 presenting another award to Frank Lewis, this time for outstanding service.

Always keenly interested in music, Frank W. Lewis was the director of the NSA Choral Group. Some of the gleemen are seen presenting a plaque to him shortly before his retirement in 1969.

On his retirement from the National Security Agency, Frank W. Lewis received a scroll signed by hundreds of his friends and fellow-employees. The scroll was later filled. Jack Gurin is on the left.

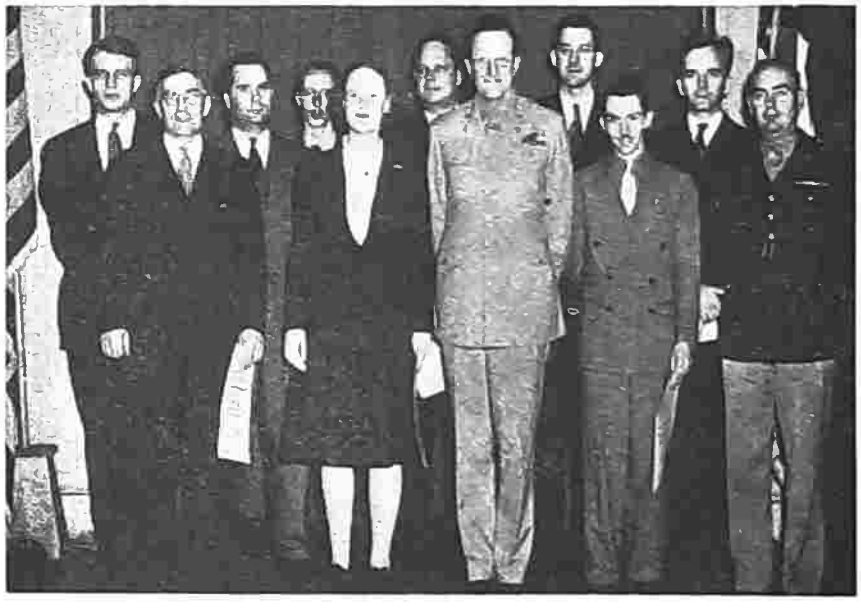
In 1949 Frank W. Lewis received an award from the likable, rough-and-ready General Ralph Canine, who was then Director of the Armed Forces Security Agency.

Here is an early picture of the NSA Choral Group, featuring some of the gleemen. In its heyday, this musical group was 30-strong.

Frank W. Lewis posed with his family before they left in 1955 for England where he was deputy to the Senior United States Liaison Officer. They served with a British Foreign Office group.

In 1969, Frank Lewis received the National Security Agency's highest honor, the Exceptional Civilian Service Award. Admiral Noel Gaylor, DIRNSA, made the presentation. Dr. Louis Tordella, D/DIRNSA, is on the right. Frank Lewis' wife, Sylvia, and two of their children, Frank Jr. and Wendy, were there to witness this presentation.

Five noted cryptanalysts of the War Department's Signal Intelligence Service gathered for this photo, from left: Art Lewis, Frank Lewis, John Hyman, Sam Snyder and Bob Ferner.



After World War II, commendations for Exceptional Civilian Service were awarded to the nine persons pictured here with Colonel Harold Hayes and Brig. General P.E. Peabody of the General Staff, who made the presentations. From left front: Mark Rhoads, Genevieve M. Feinstein, General Peabody, Frank W. Lewis and Colonel Hayes; second row: Thomas A. Waggoner, William H. Erskine, Paul V. Gerhard, Albert W. Small, Robert O. Ferner and Martin Joos.

This picture was taken shortly after World War II at Arlington Hall Station, home of the Signal Intelligence Service. The person on the left and the person cutting the cake are unidentified. Frank Rowlett is second from the left and Frank Lewis is on the right.

In this 1961 picture, Vice Admiral Laurence Frost, DIRNSA, presents another deserved award to Frank Lewis.

Frank W. Lewis and his wife, Sylvia, as they appear today. When they retired, they chose to live on the island of Montserrat in the Caribbean. Montserrat is 27 miles southwest of Antigua in the Leeward Islands.

The SIS crytanalysts occasionally took time off for a picnic. John Hurt, the Signal Intelligence Service's leading translator, is on the left. Abraham Sinkov is facing the camera.

Arthur Lewis, Frank Lewis' father, was born in 1864. Frank Lewis, on the right, was born in 1912. This picture was taken in 1929—long before Frank even thought he would become a key figure in the U.S. successes against the cryptographic systems of the Japanese military forces in World War II.

Colonels Doud, Minkoff, and Hayes face the camera at an SIS picnic.

Signal Intelligence Services picnics occasionally took place during World War II. The astute viewer can identify William F. Friedman standing next to Col. Charles B. Brown, Officer-in-Charge of "C" Branch.

This is another view of the same Signal Intelligence Service picnic.

This rare and treasurable photograph was taken by Solomon Kullback in 1940 at a picnic in Rock Creek Park, a green running from Georgetown through Washington, D.C., to Maryland. Kullback's group was an important operating section of the Signal Intelligence Service, then situated in the Munitions Building on Constitution Avenue in Washington. Some of these IDs may be in error, but they are believed to be as follows, from rear to front, Column 1: Paul W. Derthick, William Lutwiniak, George Poage, Lou Laptook, Eli Schuman, E. Stefanski; Column 2: Blanche Hatfield, Meredith Gardner, B. Bijoue, Frank W. Lewis, Solomon Kullback's son and daughter; Column 3: R. Pettengill, Paul Deafenbaugh, S. Derbyshire, Arthur Lewis; Column 4: B. Henry, D. Pettengill, James Lively, Leonard Bickwit; Column 5: Burt Lifschulz, Ann Harrington, James Lealy, Roy Johnson; Column 6: Ruth (unknown), Peggy Berthiaume, R. Coypin, R. Amberg and M. Ahlijian.

Chapter IX

TRANSPOSITION

WE HAVE seen a variety of substitution ciphers, and should at least tip our hats to the other main category of ciphers: transposition. Many solvers shy away from transpositions, not because they don't know how to solve them, but because such problems do not give quite as much opportunity for analytical reasoning. Solution often may depend upon exhaustive trials of various widths, or finding the exact method of inscription, etc., rather than a flash of inspiration as to the reason for some peculiar behavior of certain elements, or the discovery of a very important cycle, for example.

While we will deal with transposition of straightforward plain text, there are many applications of transpo as an added feature in the way of security.

In the world of communications, there are many examples of transposed elements, such as one might encounter in enciphered speech, or encryption of television signals. You can transpose a message already enciphered by some form of substitution; and the transposition of a message encoded in some way (even by a commercial code, which is hardly a "secret" document) adds a good deal to the security factor.

There are many ways of transposing a plain message (which for our present purposes will be presumed to be straightforward English literal text). The most common method is merely to write the message out (from left to right) on a prearranged width, and then prepare a transposed version by taking the columns off in some order (perhaps by a numerical key, which may be based on a literal keyword.)

Thus, as a miniature example, suppose we have decided on a format four columns wide, with a keyword of CODE. There is no necessary limit on the "depth" of the box we are to use, but practical consideration dictates that we should have no great number of rows in excess to the number of columns. A message of 50 letters might be put into a format from 6 to 8 wide and 7 to 9 deep, for instance — 6 rows and 9 columns or the other way around, or 7 x 8 in either direction.

In our sample, we would not want much more than 20 letters; my "message" is 18 letters long. Above the "box" I have written the numbers 1 4 2 3 — obtained from the prearranged CODE. To get this key, I would put a 1 under the letter which is lowest alphabetically — in this case C; a 2 would be placed under the next "earliest" letter D.

Similarly 3 goes under E and 4 fits under O (the letter last in an alphabetic arrangement of the four in the prearranged keyword.)

```
1 4 2 3
T H I S
I S T R
A N S P
S I T I
O N
```

I now copy the letters by columns, reading 1 first, 2 next, etc. The final message would probably be sent in 5-letter groups:

```
TIASO ITSTS RPIHS NIN
```

Of course the frequency count on the cipher version is exactly the same as the plain version, since we have changed only the order, not the identity of each plain value.

Completely Filled Rectangles

In the sample shown, you will note that some columns are 5 deep, while others are only 4. The rectangle was not completely filled, which would make solution slightly more difficult, as we shall see.

For our first attempt, we'll take a completely filled rectangle, and show you how to tackle the reconstruction of plain text. Suppose you have this message (appearing as Problem No. 2 in the CRYPTOGRAM's "Cipher Exchange", which ranks its entries from No. 1 to 25 more or less in terms of difficulty):

```
EOBFE OUTNE TCTTT BROPM TUHWR TISNA OHSOS UOOLA NRZEE
LEIOS NAIRE SOTYC TTDET ATIEA HMPGG LGEYF A
```

Even if they hadn't told us this was a transposition, we could be sure from the cipher frequencies that such was the case. (If someone tried to fool us by creating a cipher in which E was replaced by T, and T by R, for example, the text might resemble transposition. But the show would be given away by a repeat, or digraph cohesion, or any of the things one might expect in a substitution but not in a transposition.)

The message is 81 letters long, and while it might NOT be 9 wide, and a completely filled rectangle (which is a square in such a case) we'll try the easy approach. No crib was given, but the title "Nostalgia" might give us some clue.

Writing the message down consecutive columns, 9 deep, we get:

```
E E P S O L E E P
O T M N O E S T G
B C T A L I O A G
F T U O A O T T L
E T H H N S Y I G
O T W S R N C E E
U B R O Z A T A Y
T R T S E I T H F
N O I U E R D M A
```

Our first step is to try to place a couple of the columns that were originally next to each other, by pairing letters on the same row that have some affinity, such as TH, or NG, or (happy the day) QU.

It would be nice if by pure accident a common word, such as THE, had appeared in the original text as a repeat at the same cycle as the transposition key. In such a case, we would find two T's in one column, two H's in another, and two E's in a third, with the same number of intervening letters in each instance. (No such luck here.)

Letting our collective eye sweep across the rows, looking for letters that have specific behavior patterns, we see N and G on the second row, but nothing spectacular would result if we paired them. The CT on row 3 checks with TH and TW, but may not be enough to get excited about.

There is another H in that fifth row, however, and if we put columns 2 and 4 together, they don't look half bad. And if we add column 1 to the pair, because of the E for THE, we get a good set of trigraphs. (We could try THIS instead of THE, but the resultant text doesn't look good — AA for example.)

```
E S E
T N O
C A B
T O F
T H E
T S O
B O U
R S T
O U N
```

At this point, some of you solvers are going to lose patience, and say "when is he going to point out the obvious word in the first row." I agree, when you get all those E's and a couple of P's, you'd have a time getting any sensible answer other than PEOPLE.

And now with the ending E to guide us, we should have no problem putting the square together. As we take the columns off in proper order, we will of course put the number of the column above the square, in case we can guess a keyword.

Let's first fill in the columns that present no ambiguity:

```
    5   6 2 4 1
P E O P L E S E E
    O   E T N O
    L   I C A B
    A   O T O F
    N   S T H E
    R   N T S O
    Z   A B O U
    E   I R S T
    E   R O U N
```

When we see that L-IC AB, and recall the title, we don't even have to test the two P columns and the two E's — we merely write them in. Note the key is 3 8 5 9 6 2 4 1 7. It's not all that easy to recover the keyword which gave rise to a 9-wide key, but later on we'll discuss techniques of solution.

Incomplete Columnar Transposition

Incomplete Columnar Transposition is naturally more difficult to solve than a case in which the box is complete, since one cannot expect to find a crib or other likely words jumbled up on the same row of a writeout, the size of which is either given or is one of a few possibilities. (If you know there are 72 letters in a completely filled box, it must either be 8 x 9 or 9 x 8; 6 x 12 is possible but unlikely.) Suppose we have this text:

```
HUOIE AMYAC IITEN TLNNE UTSLI TOSFW
BCEEC LMFHO OLDMG SDLHI ELIMR CNIPF
```

If you have a crib, the burden is eased even without being given the width. If you know the width, there are ways of matching columns to get plain text. Without either of these two crutches, you may find yourself limping.

Let's suppose we are given the crib DIFFICULT, but don't know the width. If the width is less than 9, you can probably manage things — if 9 or wider, you may still have trouble determining width. Obviously if the width is 6, three columns would have digraphs in them which you must find — DU, IL and FT:

```
D I F F I C
U L T
```

If seven wide, you can expect DL and IT; if the width is eight, only the DT digraph will have to appear. We note in our message DL and IT, so we assume the width must be seven. Unfortunately there are two IT's, and we don't know which to pick, so we must try both, going up several letters from our crib, and down several letters, since we do not know whether DIFFICULT appears early or late in the message.

```
O AN
L ME
D YU
M AT
G CS
S IL
D I F F I C U
L T
H EO
I NS
E TF
L LW
I NB
```

Do you prefer SI or SL, GC or GS? Further down, however, we have a definite favorite — HE, IN, ET, LL and IN over HO, IS, EF, LW and IB. (Remember we may have to backtrack to this point, if our assumption doesn't lead to success.) Now if the box is 7 wide, and we have a 60 letter message, there must be four columns of 9 and 3 columns of 8 (36 and 24 making 60, of course.) Looking at the interval between the DL and IT (which start at position 47 and 12) we find 35, which implies we must have three sets of 9 letters and one set of 8 in between the positions of the crib. If we had been more fortunate, the distance would have been smaller and given us more useful information. However, we have enough to go on now, since we must find a T underneath a U, if the word THE appears. (It may not.) But sure enough, there is a UT to go with our columns:

```
E O A
N L M
T D Y
L M A
N G C
N S I
E D I
U L T
T H E
S I N
L E T
I L L
T I N
O M N
```

We can only choose the best stretch of 8 or 9 rows, and the top three do not look so hot; so we will confine our attention to the bottom area. Doesn't this imply that the word DIFFICULT starts no later than 5th row — possibly the 4th or 3rd? Now since the crib goes across the entire width of the box, every column must have one letter (since we have found the two which have DL and IT.) Beginning at the first of the cipher

message, we must find a letter that could fit, and it should be in the 3rd, 4th or 5th position. We have O, I and E to choose from. There is no O or E in our crib, and we have used one I, so it must be the remaining I. This fixes the general boundaries of the box, since H must be on the top row.

```
      L M
    H N G
    U N S
    O E D I
F F I C U L T
    E T H X
    A S I
    M L E
    Y I L
    A T I
```

You would have no difficulty in finishing the problem, and with the proper beginning words you can easily sort out the long and short columns to put the box in correct order. What if you had been given the width, but no crib? With a seven-wide box, given 60 letters, we must have four long and three short columns, as we have said. But which positions of the text represent short columns, and which long? It creates some ambiguity in the problem of lining up plain text. It is possible to think of the running text as a combination of letters which must be in Column 1, Column 2, etc., and those which could go either way. Since there are three short columns, the maximum length of the "uncertain area" is three. Consider the following stream, in which areas of ambiguity are marked off. (The letters outside of these zones must necessarily be in Column 1, Column 2, Column 3, etc.)

If you have any difficulty in understanding the above, think of it this way. There can be only three short columns of 8. Therefore, if the first three columns were short, the rest of the text, beginning with the 25th letter, would necessarily consist of long columns. On the other hand, if all four long columns came first, beginning with the 37th letter all the rest would have to be short. We do not really believe either of these two cases is the true one, but the areas of ambiguity are easily marked out because of the possibilities. Naturally, when we get close to the end, the areas of ambiguity get smaller, and match the beginning series in reverse — 1 2 3 3 2 1.

```
H U O I E A M Y [A]     C I I T E N T  [L N]
N E U T S L  [I T O]    S F W B C E  [E C L]
M F H O O L  [D M]      G S D L H I E  [L]
I M R C N I P F
```

You could slide these seven segments against each other, looking for good letter relationships. For instance, the common digraphs ED, TH hand NG appear when you line up the following:

```
N N E U T S L I
G S D L H I E L
```

There are only two places where you could get an E in the fifth place (or the fourth, which would also be acceptable).

```
N N E U T S L I
G S D L H I E L

H U O I E A M Y
C I I T E N T L
```

I think you'll agree the SIN and ILL look good (Even though ILY could be the end of an adverb, the double L is much better after a vowel.) If it doesn't build, OK, but let's push.

The placing of these three rows helps pin down the shape of the box, as in the method we used with the crib, so you would complete the problem either way you tackled it.

If you have no crib, and no width, you would be forced to try the last mentioned approach a couple of times, testing various widths, but would hit on the correct one, with a little trial and error.

Remember, "Patience and perSEVerance made a Bishop of his Reverence."

Before we leave the subject of Incomplete Columnar Transposition, you might consider a technique of writing out the message so that you can follow some general approach to the problem of matching columns without having to keep on asking yourself "can these two columns come together to create this digraph, and still fit the long-short requirements of the box?"

It is called "hatting" by some, since the general shape of the required matrix has a more or less curved shape like the top of a head. (To British cryptanalysts, "hatting" means randomization, reminiscent of pulling numbers out of a hat.)

The general approach is similar to the method shown of marking out the letters into areas which must be in a given column, and marking in some other way (brackets, perhaps) the letters which could "go either way". One must know the width of the matrix, of course.

The rules are as follows: Write the first column vertically to a depth of the long columns — even though in the final solution the first column may be short. Then start the second column by repeating the last letter of the first column — but commence writing at one level above the beginning point of the first, somewhat in the manner of "stairs", with the second column being one step up.

End this second column at the point where it is even with the first — which will be the point of accommodation for two long columns.

The third column is started one further step up, but this time the last TWO letters of the second column are repeated at the top of the third. (You will see this top area, above the level of the first column, is all in the "hat.") In the message we just solved, the first three columns would appear:

```
    L
  A N
H C N
U I E
O I U
I T T
E E S
A N L
M T I
Y L T
A N O
```

This "step up" approach continues for as many columns as there are "long columns" — in this case four. At that point, the "hat" levels off and starts to step down; however, since all the long columns are "used up", the writing of the last columns (equal to the number of "shorts") stops one line above the level of the "long columns." The entire box looks like this:

```
      I E
    L T C D
  A N O L M L
H C N S M G I
U I E F F S M
O I U W H D R
I T T B O L C
E E S C O H N
A N L E L I I
M T I E D E P
Y L T C M L F
A N O L
```

Now, in trying to get good digraphs, we can go by certain rules: Suppose you have a crib, or a probable word. If adjacent columns have two of the letters you are using, they must be on the same level, or one level removed, in an acceptable direction. This direction progresses upward, reading from left to right on the left-hand side of the box, but is downward on the right-hand side.

All of this sounds complicated, but if you think of a crib running across the entire width of the box, you will find that a line connecting all the letters — in this case IIUFFDC — progresses on the left side of the box by either staying on the same level or moving one step up, and on the right side by staying level or moving one step down.

Let's hope you don't have to "hat" very often, but if you do, you will find it beats worrying with each choice "can the second letter of this pair be legally reached?"

In the example shown, consider the D (in the 6th column) and ask whether it could be paired with the I in the second column just under the C (and above the I we finally chose). The answer is "yes" — but with very restricted choice of text.

The "crib line" would have to contain the O or U of the first column, the E or N of the third, the S or F of the fourth, the F or H of the fifth, and the R or C of the last — that statement being made without trying the whole combination, which is severely limited by "longs and shorts" when you actually try to write something out. If you really have a crib, such as we had, the "looking" is much easier. So much for hatting.

Route Transpositions

There are many ways in which text can be inscribed (and "outscribed"?) utilizing a rectangle or other standard shape — even a diamond or a triangle.

Without a clue to go on, solution may be just a matter of trial and error — working with logically shaped boxes (dependent on the number of letters in the message) and starting inscription in various places, with various routes.

So about all I intend to do here is mention some of the routes you may come across, if you try to solve many transpositions.

One of the simplest types is a spiral, such as:

```
ROUTETR          LVEDBYT
ARESOLA    or    ONSPOSR
SRIALVN          SAROUII
NTYBDES          ERTETTA
OITISOP          RASNOIL
```

sent as RASNO ORRTI UEIYT etc., or LOSER VNARA ESRTS etc. Notice that when you write the cipher out by columns (at the right width) you get patches of plain text going in both directions.

You can make these problems a trifle more difficult by using a numerical key to govern the order in which columns are taken out. Thus, if you had a version of the text that went TSABI TLVEO RASNO RANSP, etc., you would have to anagram the rows to find any hints, but the columns would still show things like RANSP.

Another way to write the cipher in is to start in one corner and then follow every diagonal up or down (or alternate the direction). I use numbers to show the order, with the plain being equivalent to 123456789:

```
1 2 4 7 .        1 3 4 . .        . . . . . .
3 5 8 . .   or   2 5 9 . .   or   . . . . . 7
6 9 . . .        6 8 . . .        . . . . 8 4
. . . . .        7 . . . .        . . . 9 5 2
                                  . . . 6 3 1
```

You can also have boustrophedonic pattern on either rows or columns:

```
A B C D E F
L K J I H G
M N O P Q R
X W V U T S
Z . . . . .
```

or

```
A J K T U .
B I L S V .
C H M R W .
D G N Q X .
E F O P Y Z
```

Grille Transposition

A favorite method of transposition involves the use of a cardboard or plastic grille, in which holes are cut, or squares are carefully "windowed" (at least a quarter of an inch square, and preferably a bit bigger) so that it can be placed on some sort of format, and letters either inscribed through the holes or read from the paper underneath, for decryption.

Transposition grilles are usually square — 8 x 8, 9 x 9, 12 x 12 for example. (A friend points out that more exoctic shapes such as diamonds are theoretically possible.) Exactly 25% of the area (with even dimensions, and a slightly lower percentage for odd dimensions) are "holes" through which a plain letter can be read.

For a message close to 100 letters long, a 10 x 10 square would probably be utilized. Setting the grille against a sheet of paper, being careful to mark the bounds of the square so that it can be placed in the exact position later, the sender would write the first 25 letters of his message in the holes, or open space.

The grille would then be physically turned 90 degrees clockwise, and the next 25 plain letters inscribed on the same piece of paper, with the grille set exactly on the guide lines made at the first setting.

The next 25 are filled in after the grille has been turned the same way another 90 degrees, with the remainder of the text inscribed during the final turn of the grille. It is not absolutely necessary to have the last portion of plain text the same number of letters as the first three inscriptions (25 in this case).

But it is a requisite that the grille be filled in with something, so either X's are used (if a very few fillers are required) or dummy letters (either random letters or a nonsense filler, being used to indicate the genuine end of the message by some prearranged signature or word.)

Too many X's at the end would be a grave mistake, since one could reconstruct a good deal of the grille from such insecure usage.

The arrangement of these holes is carefully planned, so that if a hole allows a plain letter to be written in at position 1 on the first setting, as the grille turns for three more inscription processes, a hole will never appear again over that first position, number 1.

Let us imagine a small 4 x 4 grille, with O to represent a hole large enough to inscribe a letter onto a sheet of paper placed underneath the grille.

```
O . . .
. . . O
. O . O
. . . .
```

Suppose our message is THE LENGTH IS RIGHT. The first four letters would be written in through the holes as they appear above:

```
T . . .
. . . H
. E . L
. . . .
```

The grille would be turns 90 degrees clockwise, and four more letters written on the same paper, causing it to appear:

```
T . . E
. N . H
. E . L
. G T .
```

Two more turns, two more inscriptions, and the final version on the piece of paper will look like this:

```
T I G E
H N I H
S E H L
T G T R
```

Grille transpositions are just the stuff agent systems are made of. They are reasonably effective if properly used (never using the same grille for two different messages.) But the tendency might exist to use a fairly small grille (say 10 x 10) for a very long message (say 485 letters) by breaking the message into four parts of 100 letters and a carry-over for the remainder, using nulls for the last 15 positions.

These "parts" of the message, because they use the same grille and are therefore inscribed (and taken off) in exactly the same arrangement, could be anagrammed — and it might even be a good idea to include such a problem later on in this teaching effort. For now we will jointly tackle a medium size grille problem, in which a crib is hidden. Although sent in five-letter groups, as usual, the message has exactly 100 letters, so the grille must be 10 x 10, and it is written out for you below at the correct width. The title given is "mathematical paradox" and the tip is OBJECTS.

O I A F U R N G Y O
A H P I S O E E R B
J S T U E M T B C A
O Y S B N C E T T T
E A S I E O P Q A F
R I N T A S U S I I
E D V O L A N X F L
I E I N N F C I A C
A L T L N T T Y E O
E T L O N Y A R P M

Every solver has his own approach to problems, and the next ACA member you meet may have an entirely different method of solution. But I will show you two ways of handling the recovery of plaintext that I happen to like.

The first may put a premium on your ability to visualize spatial relationships without standing on your head or painstakingly counting "four up and three across equals four down and three back" while juggling four separate sections of the plain text. The second is an exercise in simple arithmetical relationships, tedious but more or less foolproof.

I will show you the first method as the best thing to use as long as progress is being made. (If you get bogged down, the other method might clarify things in your mind.)

Solution involves the building up for four separate 10 x 10 boxes, each one equivalent to the position of the grille at one of the settings. Therefore, you will wind up with 25 letters of plain in each box, distributed according to the placement of the holes in the grille.

The first thing we must do is locate the word OBJECTS in the square of cipher. Remember that the holes in the grille were so made as to permit the viewing of these letters as consecutive positions, with no other windows present at that setting between the O and the S.

The J is a dead giveaway, since there is only one occurrence. And you find all the correct letters before and after, some of which are fixed, and some presenting a certain amount of ambiguity. For example, the B is fixed, as is the E, but you have a choice of two C's before a flock of T's.

Although the O closest to the B is the best bet, it is barely possible that one of the two O's in the first row is correct.

There will be an average of between 2 and 3 letters per line, which means that we could get a row with no hole in position to read a letter on that line. And some other line might have five plain values present for one of the four squares we are going to build up.

You can't rely on it completely, but there is a tendency for holes in the grille to take some pattern, such as several openings adjacent on some diagonal.

Usually there will not be many places where two or more consecutive cipher letters are taken for the plain, since the grille gets a little bit awkward to use if there is a large area of consecutive holes — for instance in a contiguous diamond shape where the center might "fall out".

We will position the letters of OBJECTS in a diagram I have labeled Box I. Since we will call Box II the setting where the grille has been rotated 90 degrees, Box III will represent one more turn — 180 degrees away from the original setting of Box I. And this Box III is the "weak link" which permits the testing of ambiguities found in the placement of the crib in Box I, so we will tackle it first, by entering the letters that would appear if the windows cut out according to the crib (in Box I) are correct.

This is where it helps to visualize Box III as a reverse mirror image of Box I. If we take the sixth letter of the second row in Box I to locate the O of OBJECTS, we must utilize the 5th letter of the ninth row in Box III as the last position of seven plain values which must make sense.

Therefore, putting in the uncertainties as little letters, and the positive values in capitals, we get for Box I:

```
o . . . . . . . . o
. . . . . o x x x B
J x x x E x x x c x
x x x x x c x t t t
x x S . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
```

This gives us at the same time values in Box III — sort of like a mirror-image:

```
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . . . .
. . . . . . . S x x
e d v x l x x x x x
x e x x x F x x x C
A x x x n
e                 m
```

If the square we started with (Box I) reads OBJECTS, Box III has a string of letters which read S(e,d,v)(le)FCA(nem), with the capitals "must" letters, and a choice of one value of the small letters in brackets. To me, the best possibility is SELFCAN.

Note I have put in x's to denote the areas which cannot have a "window".

Since the first crib letter in Box I has three possibilities, we can only guarantee the x's where I have placed them, but if we used the top line for the O, naturally we could extend this area of "no opening".

Let's put in what part of the grille we establish by using SELFCAN and rotate it through the full cycle, showing certain letters in Boxes II and IV.

We were fortunate in landing on a Q in Box II, and the only U within reach happens to be fixed in this Box. So we can say the spaces between the Q and the U have to be marked out — no window. We could go back and add the information to the grille we are building up, but it might be worth a look at the possible text with the QUI.

Require? There is no R within reach.

Box I

```
. . . . . .
. . . . . O x x x B
J x x x E x x x x x
x x x x x C x x x T
x x S . . . . . .
. . . . . .
. . . . . .
. . . . . .
. . . . . .
. . . . . .
```

Box II

```
. . . . . x x G . .
. . . . . x x x . .
. . . . . M x x . .
. . . . . x x . .
. . . . . x Q x x
x x x x x x U x I .
. . . . . x x x .
. . . . . x x x .
. . . . . x x x .
. . . . . A . P .
```

Box III

```
. . . . . .
. . . . . .
. . . . . .
. . . . . .
. . . . . .
. . . . . S x x
e d v x l x x x x x
x e x x x F x x x C
A x x x n
e                 m
```

Aha. EQUIVALENT. Let's put that in, add to the grille information, and put in as much as we can in all four boxes. Remember that a letter once used in some box or other cannot be used in any other square, and that position in the grill becomes taboo in the remaining three boxes. For that reason, I usually mark off an "x" in the corresponding position of the three boxes for every plain value I put in.

Also, just to keep things straight, I draw a line through the cipher in the original square, so that in working with the boxes below, I can immediately see whether a letter is available or not.

You will note that in Box IV, on the third row, where you seem to have a choice between M and T, in fact there is no choice, because the M was used in Box II. Moreover, since IV is the "reverse mirror" of II, you get a chance to fill in a lot of text in the places that EQUIVALENT donated to you for Box IV.

With this much text to go on, I am sure you can finish up the entire problem. The incomplete boxes below represent a knocked-off job, with no attempt to go back and follow the consequences of any obvious plain entered. But it may give you a chance to confirm your understanding of the procedure. (Why did he put an X there? What does CONTAINS in box IV do for me? Which of the four boxes starts the message?)

Box I

```
. . . . x x N x . .
. . . . . O x.x x.B
J x x . E x x x x x
x x x x x C x x x T
x x S . . x P x . .
. . . . . x x s . .
. . x . . x x ? . x
. x . . . x C x A .
. . x . . x.x Y x .
. . . . . x x x x .
```

Box II

```
. . . . . x x G . .
. . . . . x x x . x
x . . x . M x x . .
. . . . . x x . . x
. . x . . . x Q x x
x x x x x x U x I .
x x V x x A x x x L
x E . n n x x x.x x
x x T . . . x x x .
. . . . . A . P .
```

Box III

```
. x x x x . x x . .
. x P x x x . x . x
. S x U x x x . x .
x . . . . . . . . .
. . . . . . . . . .
. . . . . . . S x x
e d v x l x x x x x
x e x x x F x x x C
A x x x n . . . . .
. . x O x x . . . .
```

Box IV

```
. I x F . . x x . .
. x x x . . . E x x
x x x x x x T x C x
O x x x N x x T x x
x A x I x x x x x x
x x N x x S
```

Alternate Method of Solving Grille Transpositions

If I have trouble seeing plain text using the four boxes described above, I find the following method makes it easier to follow possible plain along a horizontal stretch, rather than continually looking up at the original matrix to see whether possible plain can be reached.

The only trouble with the method is that is a rather cumbersome chart of space relationships — showing that if you take A plain in position X, B plain must go in position Y, etc.

Suppose the grille is 10 x 10. I first write out the normal position series for an arbitrary Box I — 1,2,3,4,5 etc. Then I write out a corresponding series for Box II underneath. You can see that position 1 of Box I is position 10 of Box II, as the grille is turned. What about position 2? It becomes position 20 in Box II. What about position 11 in I? It becomes 9 in II. And so on.

Considering Box III, what number will fit under the 1 of the straightforward series of Box 1? A little thought shows that first goes to last, second to to the penultimate, etc., so that the two numbers must add to 101 — 1 and 100, 2 and 99, 3 and 98, etc.

As far is Box IV is concerned, by the time the grille has been rotated three times, I-1 will turn out to be IV-91, 2 is 81, and so on. You will note the complementary relationship between II and IV is the same as I and II, in that the 13th position of II, for example, relates to the 88th position of IV.

All four series are written out horizontally, and the cipher text is then copied underneath the string of numbers. Four more rows are left for the inscription of either a plain letter or an X. Obviously if we put an E on the second row, that position of the other three rows has to be marked out.

One effect of this method, cumbersome as it appears, is that it is somewhat easier to follow the plain running along the various levels, with all the "impossibles" x'd out. One can see at a glance places where some letter must be fitted in (since the other boxes have strings of X's between plain letters.

A fairly large area of cross-section paper is required. But since the complicated series of numbers is unique for each width, it is possible to merely use them as a "look-up" table, with your text written out under a straight numerical listing of 1-2-3 etc., referring to your table (which presumably is in the little notebook you keep handy for all crypt references).

Let's take an example of a grille — and I have prepared an 8 x 8 for you, since it won't take up so much space.

Naturally the relationship between positions is different from the one I describe above — the magic number of "totals" will be 65, for instance, instead of 101.

The messages starts out THELARG (saving you the nuisance of looking for a crib, which you could easily find.) You will see that I put in x's to mark places which cannot hold a letter; if there is a plain on one level, the other three naturally have to have x put in.

You will see the numbers above the T are 2 (because it is the second position) 16 (meaning the 16th letter is on the second level) 63 (telling us the 63rd letter is on the third level) and 49 (Box IV using the 49th letter). Similarly the H gives us more such information. I have written some maybes in small letters:

Box I	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
II	8	16	24	32	40	48	56	64	7	15	23	31	39	47	55	63	6	14
III	64	63	62	61	60	59	58	57	56	55	54	53	52	51	50	49	48	47
IV	57	49	41	33	25	17	9	1	58	50	42	34	26	18	10	2	59	51
Cipher	G	T	A	H	R	E	W	R	E	I	L	T	A	H	A	P	O	S
I		T	x	H	x	e	x	x	e	x	L	x	x	x	A	x	x	x
II	g	x		x	R						x		A			P	x	
III		x		x	x						x		x			x	x	
IV		x		x	x						x		x			x	O	

Box I	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36
I	22	30	38	46	54	62	5	13	21	29	37	45	53	61	4	12	20	28
III	46	45	44	43	42	41	40	39	38	37	36	35	34	33	32	31	30	29
IV	43	35	27	19	11	3	60	52	44	36	28	20	12	4	61	53	45	37
Cipher	T	R	M	G	H	H	E	S	I	A	E	C	T	L	R	L	F	I
I	x	R	x	G	x	x	e	s		x			t	x	x	x	x	x
II	x	x		x	H	x			i		x	C	x	L	x	x	x	x
III	x	x		x	x	x				x		x	x	x	x	L	x	x
IV	T	x		x	x	H		s			e	x	x	x	R	x	F	I

Box I	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55
II	36	44	52	60	3	11	19	27	35	43	51	59	2	10	18	26	34	42	50
III	28	27	26	25	24	23	22	21	20	19	18	17	16	15	14	13	12	11	10
IV	29	21	13	5	62	54	46	38	30	22	14	6	63	55	47	39	31	23	15
	C	I	L	C	B	E	O	R	M	R	Y	A	L	P	R	D	I	A	E
I			x	x			x					x	x		x			x	x
II			x	x			x					A	x		r			x	E
III		x	L	c	x	x	O	x	m	x	x	x	x	p	x	x	x	A	x
IV			x	x			x					x	L		x			x	x

Box I	56	57	58	59	60	61	62	63	64
II									
III									
IV									
	P	T	S	R	S	E	O	D	X
I						x		x	
II						x		x	
III				R		E		D	
IV						x		x	

You can easily read the last part of the third row as being COMPARED, so that word is put in and the necessary tidying up of the original "THE LARGEst" assumption is made. One quickly gets used to such things as finding out what the consequences are of the C at position 40 in Box III.

One could scan across the series of relationships and ask "where is 40 on the third row?" and find it at position 25 very easily, since Rows I and III are complementary (adding to 65) and the backwards order is so obvious.

It is a little trickier to follow the path after fixing a letter on Row II or IV, but you get familiar with the procedure.

You notice you could get WITH at the beginning of the fourth row. What better thing to follow COMPARED? Little letters? This gives OT to start OTHER.

Remember the best place to check Row IV is against Row II — which must have the same pattern (of consecutive letters) running backwards If WITHOTHER is correct, so is LIBRARIES.

I didn't clean this up, leaving some of the polishing to you, but you will concede that it is easier to see consecutive text written out in this manner than it may be to see what is happening within the boxes.

Perhaps I should add a word about a variation on the transpositions under discussion which is too complex for inclusion here, but which is familiar to most amateur cryptanalysts.

The "toughie" is double transposition, in which a message is enciphered by one of the methods we have just encountered, and then re-enciphered, perhaps with the same size rectangle and key, but more likely than not with a different key (at least).

This procedure defeats most of the usual attacks. It is possible, if a crib is known, to work out the complicated relationships and recover the set-up, but unless there are a few rare letters this can be a murderous job.

En passant, some people may have the misapprehension that putting two keys on a simple substitution cipher would vastly increase its security. (The double substitution consists of enciphering the message monoalphabetically with one mixed alphabet, and then reenciphering the result as a second step with an entirely different alphabet.)

The extra effort is scarcely worth the candle, since the message remains a simple substitution, even though you may have changed all the E's to R the first time around, and the R's to X the second time around. The only (and slight) increase in security comes from the fact that one cannot take advantage of alphabet reconstruction, ordinarily.

Of course, if one puts a TRANSPOSITION on top of a SUBSTITUTION, the result is a much more secure cipher — too difficult to include herein. However, special solutions (involving such things as common beginnings or endings, long repeats, and such) can make inroads into the system.

Even amateur cryppies may be acquainted with certain historical cases, such as the ADFGVX system used in WW-I by the Germans. The name derives from the coordinates of a 6 x 6 box containing letters and numbers. The plain message was enciphered digraphically — giving such groups as FX GA VV, etc., and then the resultant cipher was further treated by a single transposition key.

During that war, the Allies were able to read a certain amount of traffic when break-ins could be made by means of special solutions (as mentioned above) but a general solution was not available until Friedman and Co. worked one out much later on.[1]

NOTES ON CHAPTER IX

(1) Should the reader desire to learn more about the ADFGVX system, as well as read what the Master has to say about many other phases of Cryptography, he could do no better than invest in one of the Riverbank publication reprints.

The Aegean Park Press puts out a number of Friedman books, and the classic *Elements of Cryptanalysis* touches on everything from the analysis of ciphers to the analysis of codes. The editors have included a number of new problems, and the last one happens to involve the ADFGVX. It is especially interesting to me to find proof that certain solutions (particularly the "special" variety) are well within the range of the beginning cryptanalyst. There is also a code recovery problem in the book, and enough information is given to make solution a "fun thing" — perhaps more interesting to the novice than many of the so-called simpler systems. The address of the Aegean Park Press is P.O. Box 2837, Laguna Hills, CA 92653.

Chapter X

PROBABILITIES AND STATISTICS

BY NOW you have probably decided that I'm not much of a mathematician; you have seen few formulae, no Sigma signs, and probably the Xs and Ys were part of the text. But I have learned to cope with simple probabilities, and stress the fact that if I can do it, you can do it.

If you are like me (forget the "I" and the implied "am") you may be put off by complicated mathematical expressions and large Greek characters I see sprawled around the usual text on the subject, or even the fairly understandable article in a good encyclopedia. So let's stick as much as possible to simple English and a layman's explanation.

Complete ignorance in that area is bound to detract from your decision-making capacity in many fields — proving a mere handicap in puzzle-solving, but possibly an expensive omission if you have "a little on the corner" at bridge, or indulge in the occasional poker game.

I am deliberately going to keep this at the apples-and-oranges level, gearing it to a practical understanding of your chances of getting a certain number of "hits" on a cipher "overlap", or how much you should wager on drawing an Ace.

The Universes of 'With' and 'Without' Replacement

One of the first distinctions we are going to make is between those entities which have a fixed probability (such as odd or even in certain games of chance, or the throw of dice) and the situations which are changed by the removal of certain elements (such as the remaining cards in a deck after five cards have been exposed.)

The first part of the universe above will include situations in which you will see the words "with replacement". Obviously if you put the 26 Scrabble letters in a hat, each of which represents a different letter of the alphabet, you will have one chance in 26 of drawing an E, or a T, or whatever. But if you keep that letter out (a B, for example) the chances now are one in 25 of drawing an E, but " zilch" of drawing a B. However, if you replace the letter in the hat, you're back to the same old stand.

Perhaps more importantly, the "Without Replacement" rules that govern the play of Scrabble are likely to be somewhat complicated, but you are very likely to have to cope with them in many card and board games. You may not consciously try to remember exactly how many Es there are in a Scrabble set in pristine condition, but you will have a feeling for the relationship between number of tiles of a given letter and the point count assigned to it. Obviously there is much more chance of drawing an I than there is of drawing a J.

It would take a Music Hall Memory Marvel to quickly figure out how many of each letter remain in the bag after the board has been rather well filled, but expert players will almost certainly be able to say that they should throw away that "Q", since all the "U's" (and "blanks") are gone, and those on the board permit of no precursor.[1]

On the other hand, many games of chance offer opportunity for learning a few probability figures which do not change with each play.

Since "With Replacement" implies a constant source from which one is making (or can make) selections, it is covered by the same rules as one might use with a deck of cards when the seal hasn't been broken, or two dice, or a roulette wheel.

The distinction made with situations where there is no replacement, is important. (It is usually a slightly tougher job to keep the general mathematics in mind in the "no replacement" world, since the odds are constantly changing.)

Now let's consider a few typical problems that might occur in the fixed universe or "with replacement" category.

The Constant Universe (With Replacement)

The familiar "bones" (which reach back for a couple of millennia, at least) offer a starting point which is very easy for the beginning "probability finder" to master, since the combinations are relatively limited.

The more experienced reader (or the more proficient crap shooter) may well want to skip this and following examples until a level is reached which may have some more challenging appeal.

Since there are six sides to a cube, it is obvious that there is a one-out-of-six chance of throwing a "one", or any other of the six possible numbers, using one die. Throw two dice together, and the combined totals, varying from two to twelve, offer a slightly more complex table, if you are trying to figure out the probability for each.

How many ways can the dice turn up? 36, which is a good deal more than the 12 at which a rank novice might hazard a guess. The reason, of course, is that some numbers can be arrived at by several combinations — and more importantly, by combinations which seem the same, but are different.

There is only one way of "making" a two ("Snake-eyes") or twelve ("Boxcars"). You need a single spot on each die for 2, and a six on each for 12. But getting a three can happen if you have 1 on Cube A and 2 on Cube B, or the other way around.

Similarly, 11 has two combinations of the 36 possible which yield this number — lucky if it is a first throw.

When I tell you that 4 has 3 ways, you'll probably notice a nice little pattern of symmetry being established:

Point	No. of Ways	Involving		Point	No. of Ways	Involving	
		A	B			A	B
2	1	1	1	6	5	1	5
						5	1
3	2	1	2			2	4
		2	1			4	2
						3	3
4	3	1	3				
		3	1	7	6	1	6
		2	2			6	1
						2	5
5	4	1	4			5	2
		4	1			4	3
		2	3			3	4
		3	2				

Naturally this progression reverses itself, with 8 having 5 ways to obtain it, 9 having 4, etc.

Spending a few paragraphs on such a simple matter may seem trivial, but the exercise points out two things. One is, that quite often some rhythmic progression or a very simple formula will assist you in quickly visualizing what the odds are that a certain event will happen.

You might be surprised to find that many people not used to dealing with such matters would consider it entirely logical that since there are eleven numbers from 2 to 12, the odds should be one in eleven of throwing a 4 or a 7 or whatever.

The other statement might well be "Don't forget that the 'With Replacement' — or Constant — Universe is immutable."

The reason for this warning is that I am slightly shocked to see the following scene when I happen to attend "Vegas Night" at the local Golf Club. In the laudable attempt to raise money for repairs to the greens, or something, they set up very modest "gambling" equipment, and run a game of "21" (or "Blackjack"), a crap table and a roulette wheel, etc. (Chips are very inexpensive, and the unlucky person who drops a hundred of the local currency —worth a little more than $30 U. S. — feels he is merely contributing to the maintenance fund.)

Now comes my shocker. Very intelligent people, successful in areas of finance, the arts, government, and so on — and their spouses, who are equally savvy — watch very carefully the turn of the wheel, and if Odd has come up several times in a row, they quickly put down a lot of bottle caps (or whatever) on EVEN.

Similarly, if High has had a number of runs, they announce that it is time for LOW to pop up. If colors are involved, Red and Black get the same treatment, and you will hear conversation lauding the wisdom of making heavy bets on the relatively low-paying "two-choice" situations as opposed to seeing all of your money go down the drain on number 13.

If I butt in enough to make a mild rejoinder that the odds are exactly the same for each turn, regardless of what has happened during the last ten minutes, I get a slightly glacial put-down — I just don't know what really happens to that magic Wheel of Fortune.

(Of course if someone dropped the roulette wheel on its way to the club, and 13 came up eight times in a row, I'd be a little suspicious concerning the symmetrical balance of the darn thing.)

I know I'm going to get into trouble with many readers for my thoughts on the following matter. In some of the bridge games I enjoy at informal parties (with the same samples of Homo Sapiens referred to above) I get my wrist slapped for the bad habit I possess of picking up any cards I can see available and starting to shuffle them.

(I seem never to cut the cards or deal out of turn — only the instinctive shuffle, during which time I am perhaps talking, or reviewing the last hand.)

Anyway, what gets me in trouble with many players is the fact that, after the usual ritual of cutting for deal at a new table, I do not carefully note which deck was used for the cut, and make sure the lucky "high" person gets the equally lucky deck.

Horrors. Am I trying to cheat by depriving the dealer of the pet deck which has already displayed its favoritism? (Of course the high cutter has a perfect right to the choice of deck, and I am only lamenting the sometime reason for choice.)

This person may be the same one who has been very explicitly filling me in on the run of bad (or good luck) which has been his lot for the past three-and-a-half weeks.

While I agree that some people seem to be luckier than others (due, I claim, to the fact that they prepare the groundwork very carefully for "the lucky break") I cannot believe that from the sixth of June to the ninth of July my pal Joe, or whoever, was destined for a run of Yarboroughs, and after that period can settle back to getting hands averaging at least 15 points instead of 10 (which should be par for the course.)

In actual fact, there are many reasons for seeming runs of luck, including the fact that one is either under the weather or on top of everything physically, or that a nagging bill from the dentist may be occupying that part of the mind that should be considering a fancy bidding system.

But the most important consideration is that there WILL be periods of "bad luck" due to the purely chance factor of "number of bad hands in a row" at the extreme edge of the theoretical curve of distribution. (The other side of that coin implies there will be times in your life when you can seemingly do no wrong at the card table — for you are on the other side of the curve.)

Obviously, the more you play cards, the further out this curve will spread, in theory, since the number of incidents governing the probabilities will be so much larger. If you never match pennies, your chances for a run of 17 heads don't loom as a large factor in your life.

The main point for consideration is that you cannot possibly tell at the time what exact type of trough or crest you are riding — it could flip-flop at any moment, and go into the books as just one of the many randomly distributed "events" that you are probably destined to suffer or enjoy.

Some of the strikingly "unusual" but "within the range of probability" events occur on the charts headed "out of 50,000 occurrences —" (You should live so long).

If you look at a very few samples of reputedly "random"[2] digits written out as a long stream, and consider the odd/even character of each, you would be highly surprised to see 12 even numbers in a row, perhaps. But if you are evaluating many, many streams of this sort, you can be assured that there will be "funny" patterns of all sorts, since they will be statistically "expected."

The layman (like me) has to get used to the fact that oftentimes you are making thousands upon thousands of comparisons in a test that seemingly doesn't involve all that many figures. For instance, in a page of 5-digit cipher groups (say 20 rows of 8 groups each) we might notice a few repeats, and ask "are they significant?"

In such a page we have 800 digits, which yields close to 320,000 comparisons between individual strings of width 3,4,5 or 6 etc. — the number of 10-wide stretches is of course only a few less than the number of 2-wide possible combinations.

With 320,000 possible matches (and aren't you surprised to see that there are so many comparisons within so few digits) you would naturally expect to see a 3-digit hit one time in every 1000 comparisons, a 4-digit hit once in every 10,000 and so forth.

Is it apparent that you probably should not get very excited if you see stretches within the groups such as these: 97759 26110 84091 followed later on by 87592 61103. You would expect to find a couple of six-digit hits, and getting an 8-digit string to match isn't all that rare.

My friend would say "if you make that same test such and such times, you will expect to get a 9-digit hit so often, at least this many 5-digit hits so often" and so on. So what's the big deal?

But referring again to the club member who doesn't think I appreciate his "run of bad luck" or the reason for his switching to black after red came up ten times in a row, I'd better leave this subject before I offend some reader by admitting my lack of faith in Astrology, or admit to an enthusiastic reading of the Skeptic and an equally unenthusiastic acceptance of spoon-bending, UFO sightings, and such.

Now one of the main reasons for keeping abreast of the "odds" situation is to ascertain whether or not something is a "good bet." If the odds are "even" (and no wonder foreigners have some trouble with English!) as they should be on an even-odd call, or high-low and black-red in certain games, you should expect to pick up one "unit" as a winner for every "unit" you have bet.

Don't forget you also pick up the bet you had placed. This poses a philosophical problem for some people. After all, you didn't lose the unit you placed as a bet. Does that mean you are two units ahead, or one unit — the winnings — ahead?

Unfortunately the betting window keeps your wager at the track, so your "investment" must be subtracted from the winnings in strict accounting.

If the odds are four to one, you should get four units back for each unit laid down. But in reality, does the four-to-one odds represent the true situation in either of the two universes under discussion?

I heard a TV program last night, featuring an expert in "what happens at Vegas" and gleaned a few gems of information. Since I am not really a Vegas type, I hadn't realized some dens of so-called iniquity have varying (and possibly misleading) terminology.

I think I understand the meaning of such terms as "5 to 1" of course. For every chip you put down, the nice girl dispensing such largess puts down five chips, leaving your bet, as well, so that you pick up six chips. But in some cases, the odds may be listed as "5 FOR 1." And this is a horse painted quite differently.

If you see the word FOR instead of TO. the amount of chips left to the winner is the posted figure INCLUDING the bet chip lying there. So that 5 FOR 1 is really 4 TO 1. So you had better watch your English.

Incidentally, I find the nomenclature a trifle inconsistent, in that odds of 4 to 1 mean you win 4 for every 1 bet — IF you win. So you should really say it's a 4 to 1 shot against your winning — or there are 4 times more chances of your losing than there are of your winning. This is not the same as saying you have one chance in every 4 of winning, as can be shown by bringing the odds down to 2 to 1 (where you do not have one chance in 2 of winning, since that would be an "even" bet) and finally 1 to 1 (where patently you do not have 1 chance in 1 of winning.)

So one should really add the two numbers together to get a realistic figure for the number of events being considered. Thus, in a 4 to 1 situation, you should win one time out of 5 and lose four times. This represents an 80% chance of losing and a 20% chance of winning, which is perhaps the most accurate way to think of it.

In many areas, such as races and other sporting events, the reasoning is very subjective, and almost impossible to delineate mathematically. We will leave these areas to the Belmont attendees and the Dodger fans, and treat only those matters in which there is a definite rule for establishing the probabilities.

Naturally you can't always expect a one-to-one correspondence when gambling, matching the odds exactly, since somebody has to pick up the tab for fun-and-games.

In some areas (such as slot machines) you know that the tab is hefty, but have little chance of figuring it out. (Machines can often be set to pay however little percentage of the take they can get away with.)

In the case of roulette, what seems to be a completely even break for the player is taken care of by some small gimmick, such as the 0 and the Double-0. (Some wheels have 37 positions, some 38; all have 36 cardinal numbers.)

When these "0s" come up, nobody wins (except the house) and this seemingly small factor is enough for the establishment to stay ahead, because of the enormous amount of play over a period of time.

The house could take its chance, like the better, on ordinary play, but in actual fact the odds listed for many games may be slightly altered, or rounded out in favor of the house.

The program I mentioned above cited several cases of such, wherein true odds of, say 4 out of 33, would be posted as 1 out of 8, in effect. Doesn't seem like much, but in other instances the margin of error might mark the difference between a reasonable bet and a sucker bet.[3]

The odds are simple enough to figure out — 1 in 2 of being lucky on the black/white, the odd/even or the high/low split. (Odds posted: 1-1)

On 1-12, 13-24 and 25-36 there is a 1 in 3 chance of winning, so the odds should be 2-1. For a given number, the figure is 1 in 36, or 35-1. I am deliberately ignoring the 0 and 00, which alter these odds a trifle, since the odds on "even" should really be 18 in 38, or 9 in 19, and those for a specific number 1 in 38 instead of 1 in 36.

Smart betters do not fritter away their chips by putting three of them on No. 7, two on No. 11 and one on 13, for instance. Such prodigality cuts down enormously on your chances of winning, since the losers (in the majority) must be subtracted from your lucky take. Winning on 13 in the above example would net you only 30 chips instead of 35.

Obviously I am not a sporty type — any one of such would consider me rather chintzy, possibly timorously putting one chip down on red, or pondering whether I can afford two on the high side.

Another example of the "unchanging universe" is a deck of cards, shuffled, but before a deal has been made. But since as soon as the first card is exposed, the odds start gyrating, we will leave cards to the "other" universe — Without Replacement.

Many situations crop up in "Constant" or "With Replacement" world where it is fun to figure out probabilities of this or that happening. For example, at a large party, someone will start talking about so-and-so's birthday, and someone will holler "What a coincidence — that's mine, too."

How many people should be at the party, if you want to make a profitable bet ahead of time that two people there will have the same natal day? The person who said "What a coincidence" might make a wild guess of a couple of hundred — after all there are 365 days in the year (usually). Maybe half that number, since things like times N-minus-one divided by two always seem to intrude in any probability figure I see.

But a little thought will suggest a number such as 30 guests, to be on the safe side of a "good bet."

The reasoning is this: The birthday of each guest obviously must be compared against that of every other guest. (That's where the N x N-1 divided by 2 comes in, since if you compare A against B there is no need to repeat the operation when you start to match B against the rest.)

What number should replace X, if you want to make sure that (X times X divided by 2) is greater than 365 by a comfortable amount?

30 times 29 is 870, and half of that is 435, which gives you quite an edge when 365 is the magic number. 29 would give you 406, a comfortable margin, and even 28 would still be in your favor, with 378.

In figuring simple probabilities which involve two or more events, one has to know when to add and when to multiply, during consideration of the combination of independent chances. The thing that makes the distinction is the choice between AND and OR.

Given a new deck of cards, which has been shuffled, what is the chance of drawing a Spade? Obviously, 1 in 4, since there are 13 of a suit in a 52 card deck. (We've pulled the Jokers.)

If I ask what is the chance of drawing a four, the answer is obviously four in 52, or 1 in 13.

If I say what is the chance of drawing a spade AND a four at the same time, it is tantamount to multiplying the fractions, giving us

$$\frac{1}{4} \times \frac{1}{13} = \frac{1}{52} \quad \text{or} \quad \frac{13}{52} \times \frac{4}{52} = \frac{52}{2704} \quad \text{or} \quad \frac{1}{52}$$

which is the chance of drawing the four of spades.

Suppose we say what is the chance of drawing a spade OR a four? There are the same 1 out of 4 chances of drawing a spade, and the same 1 out of 13 of drawing a 4. But because the "spades" include one of the "fours", we have to say there are only 3 chances of drawing a four which is not of the suit picked (in this case a spade).

$$\frac{13}{52} + \frac{3}{52} = \frac{16}{104} \quad \text{or} \quad \frac{2}{13} \quad \text{or 1 in } 6\tfrac{1}{2}$$

If we did it the other way around, counting the fours first, and then 12 chances for the spades (because we have drawn the four of spades), it would amount to the same thing:

$$\frac{4}{52} + \frac{12}{52} = \frac{16}{104} \quad \text{or} \quad \frac{2}{13}$$

It may seem an anomaly that after the 4 out of 52 chances for a specific number have been mentioned, we don't use 48 for the remaining number of cards, but because we are only entitled to one draw, there is no "first" pick for either number or suit.

In many instances, where we just do not know what the "remaining population" is, we must use the unseen total as being untouched. Therefore the "With Replacement" rules pertain.

The 'With Replacement' Universe

For example, if we deal out five hands of stud poker (in which the first card goes face down, and the remaining ones face up) after the first given cards are placed on the table (face down) a kibitzer might ask himself what are the chances of his drawing an ace as the next card?

Since he knows nothing about the 5 cards face down (one of which might be an Ace) he sticks to the 4 in 52 probability. But the player who peeks at the "down" card, and sees an Ace, knows the probabilities have changed. This gets into the "Without Replacement" realm, which we will discuss later.

Before we leave the subject, is it clear that the "And"/"Multiply" rule we talked about above is behind the changing statistics for getting a single letter, a digraph, or a trigraph?

If there is one chance in 26 of a specific letter appearing, the digraph rule could be stated "what are the chances for such and such a letter to appear, AND such and such to appear next." The 1 in 26 becomes 1 in (26 x 26) or 1 in 676. Similarly the chances for getting a specific trinome (a three-digit number) is the chance of getting a single digit (1 in 10) x 10 x 10, or one in a thousand.

And now let's see if this "With Replacement" universe applies to the simple probabilities we are about to consider in Elementary Cryptanalysis.

The Index of Coincidence

In dealing with unknown text (we'll say an enciphered English message where word length is not indicated) the probability for any given cipher letter (except possibly for the first or second letter of the text) of being a specific plain letter is determined from accepted counts made for thousands of letters — roughly 13 chances in a hundred of being an E, etc.

This would be true regardless of whether the "message" is a substitution cipher or a transposition.

However, if we know our text is a substitution cipher, the blind percentage tables work for the first letter about to be "seen", but obviously the odds change for the second letter exposed, if adjacent to the first one picked.

The "universe" is no longer the "With Replacement" one we dealt with first, and the statistics are very complicated, involving digraphic cohesion, vowel-consonant relationship, and the like.

Obviously we would have to use specialized tables in such cases, which are available, but not within the purview of this chapter.

Now what if we didn't care about the specific plain of any specific position, but were just testing whether EVERY cipher letter could stand for SOME plain letter, in a one-to-one correspondence — in other words, as a monoalphabetic substitution.

Obviously the text would be "rough" — perhaps many J's (because they stand for E) and E's (because they stand for T's) and so on, with several low frequency letters, and a few blanks in our limited text.

But how do we define "rough"?

'Way back in 1920 Friedman devised a test which is a cornerstone of many attacks today. It is called the Index of Coincidence — and you will hear cryppies talking about the "I.C." of some problem messages when they are being "lined up," since the test (among other uses) measures goodness of fit between two streams that may be enciphered by the same way

Although this might get a little too deeply involved into what is more than the "Elementary," I refer you to either Dr. Sinkov's work, *Elementary Cryptanalysis* or Kahn's *The Codebreakers* for the specific formula used, since I won't bother with drawing Sigmas and such.

But connected with that test is a more fundamental measurement which has the effect of measuring the "roughness" of text, which may vary between the rather flat counts one gets in the average complex cipher stream (supposedly random) and the trough-and-peak counts one gets with monoalphabetic encipherment of straightforward text.

The test is based on the fact that (on average) if you counted 100 letters of English plain text you would come up with about 13 E's, 8 A's, one B, and so forth.

The roughness of this distribution (or the roughness of any distribution, including randomly enciphered text) can be expressed by this formula:

$$\frac{n\,(n-1) + n\,(n-1) \ldots\ldots\ldots\ldots\ldots + n(n-1)}{N\,(N-1)}$$

where the string of little n's represent the frequency of A, B, C .....Z. and the big N represents the total number of letters being counted.

In the case of random cipher, where all 26 letters are more or less likely, the number you come up with is around .038. In the distribution of standard English text, it will be around .067.

These results are called Kappa values. In languages having even rougher distribution than English (German for instance) the Kappa value will be quite a bit higher — as high as .078 for French.

You can see that it would be a nice way of telling whether you are dealing with a simple encipherment, since a count of .06 or better puts it away out in front of the random count of .03 that one gets for "junk."

(Remember it has an application for answering the question "are these two cipher messages lined up correctly so that the same key stream enciphered each of the pair?")

Just to prove the point, let us take a typical cipher message and list the frequency of cipher values:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
5	10	1	9	0	3	5	0	5	4	0	1	8	3	0	4	0	9	10	2	6	10	2	0	0	0

Now add the (n-1) values — 20 90 72 6 20 20 12 56 6 12 72 90 2 30; 90 and 2 (naturally the 0's and 1's add nothing to our count.)

The sum is 600. We have 97 values, so N(N-1)= 9312; and 600 divided by 9312 = .064, which you will admit is rather close to the .067 we might expect for our idealized plain text.

Now take a count from some encipherment which doesn't look nearly as rough, but still may not be random.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
7	9	5	8	4	7	10	0	5	6	5	8	8	5	6	3	6	3	13	7	1	3	8	4	4	5

You get a count of 912 from the sum of the n(n-1), and there are 150 letters, giving 22350 for the N(N-1).

Divide 912 by 22350 and you have a Kappa value for the cipher of .041. This is a little higher than random, and may indicate a limited number of alphabets, such as one would get with a width of, say, 6 to 9 on a polyalphabetic substitution.

It's a little uncertain to pick out the exact number of alphabets, but statistics (if properly used) will perhaps indicate a likely number to start with. (I always used to holler for my mathematician friend to steer me to the best bet — although there should be other means of determining a short cycle, as we have seen.)

Perhaps we have talked enough about how to figure very simple probabilities so that we can put our knowledge to use in "diagnosing the unknown" in the next chapter.

NOTES ON CHAPTER X

(1) Since there are many books on Scrabble strategy, written by more professional players than me, there is no point to a long discussion of tactics. But certain "winning ways" are at least related to the subject of probabilities we have been discussing, and a few thoughts might not be amiss.

You should perhaps develop your own "table of expectancies" which would include such things as the following:

How much should you expect to get in return for the play of an "S"? Since you can almost always tack a plural word in your "hand" onto a word dangling some place on the board, you almost certainly have a "two-way count"; so why not be more selective than the poor player who settles for a count of 10 or so?

I like to hope for at least 20 points on the play above the value of the word itself, or else I will hang on to the tile for a better opportunity. What is an average "good" score for you? In order to figure this we should know the number of letters and their value.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
	9	2	2	4	12	2	3	2	9	1	1	4	2	6	8	2	1	6	4	6	4	2	2	1	2	1
Val	1	3	3	2	1	4	2	4	1	8	5	1	3	1	1	3	10	1	1	1	1	4	4	8	4	10
	9	6	6	8	12	8	6	8	9	8	5	4	6	6	8	6	10	6	4	6	4	8	8	8	8	10

Total 193

Using rough round-offs (since we are not building a bridge, or sending up Spacelab) this gives a little less than 2 points for an average tile. But there are 25 double-value markings (counting the star in the middle) and 12 triples for letters, which brings the average value to around 2.35.

In addition there are 16 double-word and 8 triple-word squares. By my own crude manipulations, I brought the average value for a letter to about 3.7 in a 5-letter word. The figure varies according to the length of the word, since there are more chances of a long word crossing the bonus areas. (We can all dream about the 8-letter word that includes two tripling values, plus a double on that Q or X you have been keeping, as well as the bonus for using all seven letters.)

Of course "averages" mean very little in a game that gives you seven 1-point tiles on the first round (and you COULD figure out the chances of that happening.) You make the most out of what you have. What are the chances of drawing a vowel? In the "With Replacement" universe there are 40 vowels (including Y) in a Scrabble set.

Since one hopes the drawing of these has been made in a random fashion, you can (up to a certain point) use a rough figure of 4 chances out of 10 of getting that vowel on a single draw.

But how many tiles are you entitled to? If you get three draws, you are on the statistically "happy" side; remember that you add the probabilities in this case, and 120 is larger than 100.

Naturally, if the game has progressed to a point where you can fairly easily estimate the number of vowels "gone", and can figure out how many letters are in the bag, you can change your odds — if you are particularly hep in the counting division.

But it ain't all that easy to count the board, add it to the "hands" (seven times the number of players, assuming you aren't close to the end of play) and subtract from 100 — all this just to figure out how many in the bag. Better use your "With Replacement" probabilities, even though they are pretty rough figures at this point.

The same type of "how much is it worth" question arises with J, X, Q, etc., as exists with the S — but this time the placing on Triple Value, combined with Double Word Value may become your fond hope. The situation is too complex to talk about, so all one can say is "Make the most of it.".

Don't waste an "S" or a "blank," look carefully around before placing a high-value letter on an "ordinary" spot, and make sure you have found the highest value word to tack onto, if there is more than one possibility.

It is to be hoped that the expertise you developed in anagramming when you were pursuing the foxy trail of the Cryptic Crossword (perhaps even *The Nation* puzzle) will stand you in good stead during a game of Scrabble.

A hint in technique for anagramming. Experts do not need any such crutch, and just look at the letters (rearranging them in their minds, perhaps). But I notice a lot of people feel they have to actually see the possibilities by either writing them down, or madly rearranging Scrabble tiles until they recognize something.

If you ARE going to write the letters of an anagram down, I suggest you put them into a wheel-like form where you can see many combinations at once.

Given N I T L E S S:

```
    N
S       L
    E
I       T
    S
```

Even though this is ridiculously simple (I'm sure you saw several answers) you will agree that it is easier to see in this form than when you are looking at a string of letters. (If you didn't see SILENTS it may be due to the fact that we occasionally forget unusual meanings with peculiar plurals. You probably noticed three others, however.)

(2) Sometimes it seems that "random" seldom is. I have had occasion to study many streams of what has been labeled as random, such as digits emanating from the keyboards of several typists who have been given the task of banging out page after page of 5-digit groups, or 5-letter groups.

One could spend a chapter showing how human frailty almost precludes a result that is even close to random, no matter how conscientious the typist may be. (Of course you could approximate random by conscious effort of checking off the number of 1s, 2s, 3s, etc. that you have put down, but that would slow down the procedure enormously and probably defeat the purpose of the exercise, since what phenomena of distribution was removed would be replaced by a new phenomenon of relationship. The distinction implied will be discussed more fully in the chapter on "phenomena and diagnostics.")

As a brief comment here, one might say that if a person tries to prepare a "random" series of digits, the typist or clerk is almost certain to have aversions or favorites, and the distribution will be far from statistically flat. If typing, a person is likely to have a left-right rhythm with either numbers or letters.

In numbers especially, even apart from the L-R factor,the relationships between consecutive digits (and "skip" pairs, as well) are very peculiar when one examines specimens of "random" obtained by either typing or hand-producing. You will see more of this when we talk about "deltas" — usually the difference between consecutive elements of a stream. (You can speak of "skip delta" as well, naturally.)

Perhaps the single most striking thing about such prepared material is the "delta zero" tabulation — you can examine whole pages and never find such digraphs as BB or dinomes such as 99 or 11 WITHIN GROUPS. Across groups, the count may reflect either a normal amount of repeat-values, or even an excess amount of such.

Typical groups one gets from typing "at random" might be in the following which , admittedly, reflects all the evils we have been talking about:

94857 38476 19283 88557 11186 38476 01928 83920 74839 73829

You notice that when I did remember to put in doublets, I overdid it. As a result of all this, we talk about "human random" (meaning it probably isn't random). Naturally there are many sophisticated methods of producing a stream of either letters or figures which is "guaranteed" to be truly random, but the subject is not an appropriate one for our present writing.

(3) Perhaps I should not have the temerity to question the knowledge or skill of an expert in gambling (which I am not) But the program I referred to had a pitchman who got off on the wrong foot, as far as I am concerned, in one of his first remarks.

Listing the three things that make the difference between a winner and a sucker, he chose one criterion that bothers me. It was his insistence that a good gambler must case the joint, on "21" or Blackjack, for example, by standing at the table for quite a few minutes, watching the value of the dealer's "up" card. If it were a consistently "good" card (say a 9 or better) — avoid that dealer like the plague. Wander around until you find a dealer whose counted average was on the low side, for the up card.

Perhaps I missed something, but the clear implication seemed to be that the "expert" strongly believes in "runs of luck." So do I, but with a difference. As we said, there WILL be good and bad nights, and there WILL be runs of good and bad "luck." The spread of these over a lifetime could possibly be predicted as to approxmate number, but are certainly not predictable as to timing.

Just as you seem to hit the top of a crest of good fortune might well be the occasion for wondering "how long is this likely to last?" The statement that this dealer is lucky, and you should avoid her, while this girl is obviously having a bad night might be understandable if you are talking about the boyfriend (or lack of one) who is meeting her after work. But blaming her for your winning or losing (aside from any speculation as to whether she may be a card manipulator) is tantamount to blaming the messenger for the news.

Chapter XI

SOLVING SQUARES AND ALPHABETS
(Square Recovery)

ALTHOUGH we have been using squares to encipher Playfair, Bifid, and other digraphic systems, we haven't really tackled the problem of how to solve them.

Our present task should therefore be to look at a few incomplete squares, and see if we recognize the method of construction on the basis of a few critical letters, and get a little practice in putting the pieces together.

Referring to Note 1 on Chapter VI, you will see several examples of the sort of enciphering square you might come across. Basically, you will see they are Rows (either straightforward or boustrophedonic), by Columns (regular or down-up-down), spiral (beginning with any corner, in either direction) or taken off the diagonals (starting any corner, with either left, right, or alternating progression).

This gives us quite a variety, some of which exhibit the same phenomena, looking at a few recovered letters.

What do we look for? One of the best guideposts is the letter Z. Of course it MAY be in the keyword (and even then, a couple of letters may tell you something) but if it is the last letter of the reconstruction order, two letters on the same row or column may tell you a lot — and three letters practically set the stage.

In most of the cases within the categories to be discussed (straightforward, boustrophedonic, spiral or diagonal), Z will be a "patterned" number — usually 55 but often 33.

Unfortunately (since it is a rare plain value) Z is oftentimes not one of the few letters assigned values on the preliminary go-arounds, and one can only hope its cipher value pairs with some good plain entities, to bring it into the fold.

It is probably axiomatic that the values which pin the order down best are almost invariably poor plain bets — J, Q, and Z being typical.

However, with experience, it is amazing how much you can build on a foundation provided by a relatively few assigned values. (Remember, we have suggested that in assigning values for any bifid system, it pays to locate that collection of letters which has most values on one row or column.

If you tackle E merely because it seems common, you might find that it is associated with only two values which repeat often, so that you end up with 12 and 13, but then have to assign the rest without the assurance that they are not 1, 2, or 3.

Of course there is an easy way around this. The limit of five columns or rows is not a law while you are in the early stages of recovery, and you may have to assign 6, 7, 8

etc., knowing that inevitably some of the numbers are really identical — if you only knew.

(In 6 x 6 squares, where there are 36 possible pairings instead of 25, you may be forced to label values of your crib using letters after you have used up the 10 digits, and eventually prove that B8 can be called 24.)

To prove the point that you can sometimes tell a lot on the basis of very little information, take any two letters from each of the eight boxes of our examples on Note 1 of Chapter VI, that happen to fall on the same row or column as Z. In order to make a sensible classification, let's "type" the boxes we described before.

Naturally, for these purposes of example, all of them must have a keyword of some sort (or be completely normal) — a perfectly random square would defeat any reconstruction short of complete solution of the problem.

All examples will list only one direction, Down or Across in straightforward row by row or column by column cases. Similarly, only the two basic spirals (starting in any corner or the middle) deserve a category.

Other types with ambiguous directions which may be indistinguishable and operationally identical will only be listed once — since it often makes no difference whether one considers a boustrophedon movement as starting up or starting down, or starting either left or right. In diagonals it matters not whether the second letter is taken from the first row or the first column.

There are systems which use several squares in conjunction with each other, where the specific direction of progression within the first and third squares are dependent on the make-up of the second.

Naturally in this case it does make a difference, but the single use of a single square is usually a no-matter-which situation.

Only the convention of naming row before column makes the A and B types (below) distinguishable, since by reading dinomes in the Column-Row order, one can, in effect, turn a square 90 degrees, and make it work.

A By Row:	**(1) L-R**
	(2) boustrophedonic
B By Columns:	**(1) Down (or Up)**
	(2) Alternating Down and Up
C Spirals:	**(1) From Corner**
	(2) From Center (or approximate center in case of 6 x 6)
D Diagonals:	**(1) From any corner, one letter, then**
	(2) on the diagnonal, then 3, etc.

E Others: **Uses of other inscriptions, such as starting on the complete diagonal, then to the top, then the opposite diagonal, etc., in which "filled in" areas are skipped when one crosses them in complicated movements.**

Now for our test on the squares we listed previously. Suppose we have a partial square, relatively constructed:

```
  1 2 3 4 5
1 Z   L
2       X B
3 E T
4   P M
5     D G
```

Remember, we do not know at this point whether we have five unique (partial) rows or columns, since the number 5 which we assigned arbitrarily might really be the same as 3, already used. (This seems to be a stumbling block for some people, since I have occasionally heard "This should be 3, but I've already called it 5 — isn't that a conflict?"

If you have evidence for such, and no conflicts when you rename all your 5's to be 3's, you are doing exactly what you must do to boil down the square to its proper proportions.).

In that regard, ask yourself these questions: Can 2, 3, 4 or 5 really be the same as 1? In the square above, the answer is "No." (ET, EZ, ML and DL knock these possibilities out.)

Can 3, 4 or 5 be 2? Answer: 4 can, but 5 cannot (GX) and 3 cannot (MP). Can 4 or 5 be 3? 4 can't (PT), but 5 can. 5 obviously cannot be 4 (XB or MD). So keep in mind that 4 can be 2 and 5 can be 3.

Now let's examine the four possibilities, A to D, which we listed above as types of squares. ZL knocks out A-1 and A-2 unless Z is in the keyword, but XB is just as bad. Therefore A is unlikely.

Regarding B types, the XG and ZE imply both X and Z in keyword. Possible, but improbable. (We can always come back and try to make a square of this type, if all else fails.)

How about spirals-C type? Several things are in its favor. L is on the same Row as Z, but M is in the same column as L — perfect for the half-way point of the "carry-over" letters climbing up the left side of the square. And P is on the same row (either next to M or with MNOP as a sequence on the second row, which is often recognizable as a cluster from the "high middle" of the alphabet.)

```
. . E . .
M p . p .
L . Z . .
. . . . .
. . . . .
```

In either case, E and T must be on the first row (as part of the keyword and D and G on the bottom row — or the other way around. Which would you pick, remembering that if D and G are in the keyword, it must be a long one, since T must also be in it. (Of course, 10 letters, the minimum to permit this, is possible.)

But lets first try E and T in the keyword But if we put T in the second position, then D must be in that column, but X poses a problem — it must also fit in the second column (to allow for XYZ) bringing G with it. (D and G obviously can't occupy the same spot.)

```
. T E. . .
  M P . . .
  L . Z . .
    X
     ?
```

So T must be in the fourth column:

```
  . . E T .
M N O P .
L . Z . .
. X     B
. G F D C
```

You can see how filling in the several new values would give you an opportunity to go back and recover more plain (in order to get more values into the square, in order to get more plain — and so on.) Try another case. Given the partial (and relative) square:

```
  1 2 3 4 5
1 . . L . N
2 . B . . .
3 . . G W .
4 M T . . .
5 O . . Y P
```

3 can be 2, and 5 can be 2, for what it's worth. I can almost get a few results by working on a spiral, but inevitably get a conflict. But notice there are four columns that could be in a type B square.

```
     G   W
T M  L N
  O    P Y
B
```

But the GL would have to be adjacent if the square were B-1, leaving some peculiar letters in the keyword — H I K, which is asking a lot. Perhaps G and L are in the keyword, but that makes for a very long keyword, since either MO or NP must be in it too.

What if the GL and either the MO or the NP were going in opposite directions?

```
. . . . .
    G   W              G   W
        x                  x
P     O Y     or  O      P Y
N T L M Z                L N Z
  B                  M T
                       B
```

You can make a tentative approach on either of the above, and would have to try things like the HIK and QRSUV fill-ins to test your assumptions. Notice you have a limited number of square possibilities?

Assuming the best thing first, that I is in the keyword (instead of H or K) you have:

```
    .  su  uv              su uv
    G  rs  W            G  rs W
    H  q   X     or     H  q  X
P   K  O   Y         O  K  P  Y
N T L  M   Z            L  N  Z
  B                  M T
                       B
```

You have a limited number of letters to fit into the keyword, and a couple of guesses would pin down a lot of shaky values. It shouldn't be much trouble to prove the E and the R, for instance, which would knock out the first square (too many letters on the bottom row) and make the final version:

```
  b   U V
E b G S W
m t H Q X
O   K P Y    A C D F I still to place
  R L N Z
```

with the MT on the third line, and the b limited to two places after the keyword, both of which assumptions make the best sense for trials. This much of a square will

probably "break" the message, if correct. If it does not pan out, you must go back to the "alternate" assumption and beat it to death.

But I'm sure you would have seen DEMOCRATIC and gone on to the next case.

What if you were given:

```
  1 2 3 4 5
1 E     I
2 R
3   L     S
4   M O   T
5 V U
  C G
```

Note that C and G have column designation, but no row.

In turn, can this be A, B, C or D type square? Conceivably 4 is 4 and 5 is 5 on a true square, and 2 is 1.

As far as the B type goes, perhaps G and I are in a column and S and T in the next to last — but hold it. Out of UVWXYZ to fill in the last row, the U and V are used elsewhere — and we can't have a column with four letters.

Perhaps ST start the last column, with three of the WXYZ letters to fill the gap. But O is on the same row as T (presumably the second row), and we have P and Q as the only possibilities to fill three places.).

Do we see any evidence of spirals? The things to look for are any of the VWXYZ positions which look good, the "climbing" left hand column which usually has four letters of the low to middle alphabet in relative sequence, and the second row, which has four fairly high letters in relative order.

Unfortunately, we have none of the high letters except UV, and the only "climbers" for the first column are LM (making the key-word begin with U and necessitating a V in the keyword.

(You can see that ST could not be in the fourth column leading to the U, because of LM.)

Altogether, it seems as though the A type of regular row-by-row keyword square is the only possibility.

```
  2 1 3 4 5
2 . R . .
1 . E   I
3 L       S
4 M N O   T
5 U V
  C   G
```

The bottom two rows look good, with P or Q between O and T, and room for three of the foursome WXYZ. But the third row certainly looks funny, and whatever happened to the first part of the carryover from A to K?

Now you see the importance of keeping in mind (on the basis of the original partially constructed square) that 3 CAN be 1, and 4 CAN be 2. It wouldn't help to tinker with 4 now, since the placement of N conflicts with E, but 3 can still be the same as 1, and we now have a square:

```
2 1 3 4 5
  2   R
  1 L E   I S
  3   C
  4 M O P Q T
  5 U V
        G
```

with G possibly 34 (unless it is in the keyword.) I suppose you recognize what you have recovered.

Alphabet Recovery

We have seen that certain uses permit alphabet recovery more easily than others. Obviously "recovery" is not the word when we are dealing with normal alphabets (in either straight or reverse order). However, "discovery" may be a better operative word in such cases, and it is possible for the inexperienced analyst to go quite a way down the road of solution before realizing ATTSTF may be the watchword, if you recall Friedman's dictum.

Recovery of one mixed sequence sliding against a normal is not too difficult if you have a few critical letters. In the very process of solution you may decide that such and such is the logical place for a crib because of its fulfilling the job of placing letters along the logical sequence of a K-1 or K-2 set-up — solution of the text and the alphabets go hand in hand.

In these more simple cases, remember that the "unusual" letters (which you must search for in cipher rather than plain) quite often give you more info about the alphabets than the juicy E's and T's which were so helpful at first.

The tail end of the "carry-over" portion, ending possibly with some or all of the VWXYZ sequence, may be easy to spot, as is the weak JK area. And it is unnecessary to belabor such points as "there is seldom a Q, in a word, without an adjacent U."

Sophisticated "automatic" solutions invariably exploit the strong tendency towards digraphic cohesion. The specifics may differ, but the statistics of such are a powerful aid in every language.

You have seen the techniques of chaining assist in the recover of K-3 usage, where a mixed alphabet is slid against itself. And you are now familiar with the the more difficult problem of solving two mixed alphabets in K-4 usage.

In all of the examples we have seen so far, the "Keyword Alphabet" has been obtained by the straightforward method of stating the Keyword (eliminating repeated letters if necessary) and following it with the rest of the letters of the alphabet (those not used in the keyword) in normal order.

However, more sophisticated techniques of alphabet construction exist. For example, the "carry-over letters" do not have to be dangled after the keyword — by convention, you could split this section in half, and alternate the placement of the letters between the segments before and after the keyword:

A C F I K P S W Y G O V E R N M T B D H J L Q U X Z

Or one or the other of these portions could be reversed — particularly the first part, since it then alternates the carryover in a logical progression on either side of the keyword:

Y W S P K I F C A G O V E R N M T B D H J L Q U X Z

A common method of deriving an alphabet is to write in a keyword on a specific (prearranged) width — such as 9 and then treat the block as a transposition. A keyword (possibly the same keyword used as a basis for the alphabet) controls the order of column extraction:

3 4 5 6 2 7 1 9 8
G O V E R N M T A
B C D F H I J K L
P Q S U W X Y Z

Using a transposition key based on DEMOCRATS (a very convenient word for a 9-wide key, to go along with the 11-letter REPUBLICANS) the final cipher is:

M J Y R H W G B P O C Q V D S E F U N I X A L T K Z

Just to test your understanding, could you recover the basis for this mixed sequence — which you may have placed in exact order against a normal alphabet, so that you know you do not have a decimation:

A P E F X G B S L Q M J N I Z O C U R H Y T K V D W

Hint: Notice such relations as might be built up if you write down what is associated with the V W X Y Z area:

```
Y K P U J
T V E R N
K D F H I
V W X Y Z
D A G T O
W P B K C
A E S V U
P F L D R
```

If the V to Z sequence (or the greater part of it) were at the end of any "box" of transposition, the letters immediately above might be very significant.

I have written the subsequent letters, also, in case the keyword is written in as a spiral, or runs from bottom to top. Remember that not ALL of the critical letters may be in the "carry-over" — one or even two might be in the keyword.

Speaking of keywords, it is often a very difficult job to prove that a certain keyword was responsible for a numerical key unless that key is rather wide. Up to 10 wide may be a formidable task — and it isn't all that simple to recover a fairly wide key — 15 or 16. You may come up with a number of possibilities, all of which give the same numerical equivalent.

The point is rather academic, since unless you can discover some specific source for the keys, or figure out a rota (the successive uses of predetermined procedures) you will still have to solve everything the hard way. The actual determination of the literal base for numerical key may be an interesting exercise, but of doubtful practical value.[1]

However, just to exercise your potential skill at such, let's look at a couple of examples. Suppose you are given the key:

21 13 10 18 6 5 16 19 1 20 7 17 14 9 2 12 8 15 11 4 3

This is long enough for solution — and because of a "favorable beginning" you may quickly get the phrase on which it is based.

My strict mathematician friends don't approve of some of my attempts at pseudo-statistical techniques; but please bear with me while I try out a method I sometimes have found successful in deciding whether or not the key we are looking at is actually derived from a literal base.

As you know, the procedure involves taking the lowest letter (perhaps an A?) and calling it number 1, then the next for 2, etc. When there are several letters repeated, you take care of the repeats before stepping to the next highest letter — for example if there are several A's, the first is 1, the next 2, and the next 3, etc., before going on to B or C.

This sets up the possibility of determining a left-to-right sweep of progression which is not completely random, for in addition to the repeats, there is often a continuing of the left-to-right progression merely by chance occurrence of a C AFTER you have numbered the A's, and perhaps even a D at the end, after the C.

In theory, out of a 21 wide stretch, we should expect about 10 "levels of progression" by chance — when we had to move back to the beginning to start the next highest sweep.

We could probably figure out how many times there would be cycles of various lengths — one digit before moving left, two digits before having to return, etc. And you could work out a chart for any given position, showing what the chances were of the next sequent digit being to the left or the right of the place being considered.

But this is too involved for our simple test. Let's find out how many cycles we have, and how long they are.

```
21 13 10 18 6 5 16 19 1 20 7 17 14 9 2 12 8 15 11 4 3
                      1             2              3
                                                   4
                   5
                6              7              8
                                           9
          10                                          11
                                                  12
       13                                 14          15
                   16                 17
             18          19    20
21
```

Well, there are 10 "turn-backs" which is just about what you would expect by random. But looking at the lengths of cycles, there seems to be more lengths of 3 than one might expect by chance. There actually are 5 cycles of 1, 2 of 2 and 4 of 3-wide, but we are pretty shaky ground if we try and prove anything on such slim evidence

However, again we will say IF the situation were what we are hoping for, there probably WOULD be fairly long cycles at those areas of the alphabet where there may be repeats — such as E and T. The fact that the number of turnbacks isn't exciting should not discourage us, since I mentioned this merely as another one of those maybe-worthless maybe-helpful tests you should be aware of.

Also, it's the old story I mentioned before — the message may not be Playfair, but IF it were, it would have all these exact properties — no J's, digraphic roughness, with no XX patterns on the digraphic cut, etc. Perhaps you can only find out by reading the message.[2]

Also, in "real life" such tests might not be run against a few keys you happened to recover — but suppose you had hundreds of cases and found a tiny "something funny" in each, which when measured in "volume" turned out to be highly significant.

Suppose you are looking for something with a "three-Sigma" significance, but get discouraged when every one of the "hundreds" mentioned above fail to hit that level when a certain test is made.

I don't know what my fussy friend would say, but if every one of a hundred tests had a SLIGHTLY significant skew in my favor, I'd be interested enough to pester him for a further test which would add up all the little "maybes" in hopes of finding a sure thing.

The numbers we wrote above are of some help in the next step, however.

Obviously 1 can be A, but may be somewhat higher. The 21 must be at least K, since we have 11 levels. Let's write down the entire range of possibilities — almost — by cheating a little and eliminating very unlikely situations. We can even use capital letters for the "best bets". Let's first jot down a reasonable range.

1-3 probably A to D; 4 and 5 C or D, conceivably E; 6-8 D/E and perhaps a bit higher; 9 anything from E up through I; 10-11 H up to O; 12 L to P; 13-15 M to R; 16-17 N to S; 18-20 O to U; and 21 conceivably T but probably U, V, W or Y.

21	13	10	18	6	5	16	19	1	20	7	17	14	9	2	12	8	15	11	4	3
t	m	h	o	d	c	n	o	A	o	d	n	m	e	a	l	d	m	h	c	a
U	N	i	p	E	d	o	p	b	p	E	o	n	f	b	m	E	n	i	d	c
v	O	k	r	f		p	r	c	r		p	o	g	c	n	f	o	l	e	d
w	p	l	s	g		r	s		s		R	p	h		o	g	p	m		
y	r	m	t			s	t		t		s	r	i		p	h	q	n		
			u				u		u								r	o		

You can probably see the phrase. Notice that I was rather casual about the above test — it doesn't pay to tackle such "iffy" problems as though you were measuring the stress of a bridge. You may see something, and you may not. Notice the nice ending possible — ED and some good values in front of it, such as N.

Remember. however, that if you take this, you must rethink your other values — 5 must be at least F, 6 at least G, etc., and you find you are limited to the bottom area of the rough chart we made above.

As a matter of fact, any choice at the beginning or ending of a range of letters — such as starting this phrase with T, or calling 16 N — very severely limits the rest of the sequence.

We might have first thought about trying THIS as a beginning, but you immediately get into trouble. If 13 is H and 10 I, what are you going to do about 12, which is in a cycle that intervenes?

As we said before, you probably won't get much joy out of trying to solve shorter keys, but at least you know how to tackle the problem.

NOTES ON CHAPTER 11

(1) I'm sure I shouldn't be criticized by the censors who worry about such things when I underline how important it is for any cryptographic system not to be subjected to the undue pressure of incorrect use. You should NOT use a one-time pad as though it weren't. You should NOT shuffle your keys around, and then send them out to different communicants next year. You should NOT (even in the heat of battle) abuse a system having rules which virtually preclude "depth of key." And if you are introducing a machine or a new system or a replacement for something condsidered insecure, make your tests "off the air" — and start out only when you re ready to live with the results.

Out of the many "successes" I know of, a good number of them were solved because of (1) Testing over periods during which practice messages (or the same exercise message) were sent over and over, in the careless manner of "what does it matter? — this text is unimportant," or (2) Tests involving the normal alphabetical order or keyboard sequences or the 1-to-10 numerical sequence.

There will naturally be some mistakes made in early usage of a system, but don't give your opponents any extra help by careless resends, talking about your troubles "en clair" and so forth. And don't strain the system by overloading it with bureaucratic nonsense. Many of you know the successes that Allied analysts had with such as ENIGMA (and other systems) due in part to the exploitation of faithfully "on schedule" messages saying, "I have nothing to report."

Christmas and New Years may be nice times to remember your friends at the other end of your communication link, but the probable content of a raft of similar-length messages circling the globe on January 1 — using the otherwise carefully guarded machine — may provide just the cribs an alert solver is looking for.

(2) I once was working on an unknown system which had a few of the earmarks of Playfair. I have said such "relatively easy" problems can be tough if you have no crib — and might not even be completely certain about the language used — whether it might not be code, or something.

Roughness at beginnings was evidenced — no big deal, but possibly the equivalent of STRENGST GEHEIM or Part One (name your language) or whatever. Among the repeats at the beginning, we noted MB NU XD a couple of times. (I have changed the real-life words to keep our security worriers happy.)

Could our beginning be NUMBER in some system? Aha. Note that IF the first digraph MB turned out to be NU, our second digraph NU is inevitably plain MB — IN PLAYFAIR. This by itself doesn't prove it IS Playfair, but the arrow-pointing is so strong that one gets enough of a moral boost to go ahead and solve the entire problem.

Chapter XII

DIAGNOSTICS

ISN'T that funny? How often have you asked this rhetorical question? And I'll warrant that very seldom was there any implication of risibility.

Although certain elements of this chapter have a bearing on other problems discussed in the course of what we have been reading, it is mainly applicable to those "word puzzles" associated with cryptanalysis.

For "something funny" is often the undignified name for what the cryptanalyst spends a good deal of time searching for — cryptographic phenomena. It may be an unusual distribution — such as finding that none of the 489 messages you are looking at start with the cipher letter "S". Or it might be a peculiar relationship between certain positions of a digital message, in which (out of many hundreds of messages) there was never a repeat between the first and last digits.

The general topic of "diagnosis" also might cover the initial phase of certain other solutions, for in many game situations, puzzles, logic problems, riddles — and possibly in more mundane situations of daily living — the most important first question that can be asked is "What IS the problem?"

And a close second is "How do I pick it up?" — in other words, what should I do?[1]

But for now, let's talk about solution of cryptanalytic problems. The task of the analyst narrows down to four main phases: (1) Find the phenomenon (or phenomena); (2) measure it; (3) explain it; and (4) exploit it.[2]

The first job may involve powerful machines and elaborate testing procedures. Alternately it may depend on the practiced eye of an experienced cryptanalyst, or as is often the case, it can hinge on the accidental spotting of "something funny" by an enthusiastic novice, who may not have "the foggiest" of how to explain the finding.

It is implicit that in the case of the machines to which we have alluded, a lot of experience and cryptanalytic knowledge went into the planning of the testing procedures. For more often than not, a complicated battery of diagnostic tests will merely be doing what many a cryppie has wished he could do if he had a thousand hands and all the time in the world.

The individual steps to be taken are probably not new — the main difference is that certain procedures which were impossible "by hand" suddenly become trivial jobs when the power of the chip is added to the diagnostic brilliance of a Friedman.

Following this "finding" procedure, whether it be done by electronic monsters or stumbled upon by a conscientious logging clerk,[3] there must be a way of measuring the "rareness" of the phenomenon before we proceed. It is scarcely sufficient for a traffic examiner to come running up and say "this distribution is kinda rough." or "I think I see too many doublets."

In the case of computer analysis, testing procedures are frequently built into the "finding" program, and levels of acceptance pre-set. The folly of printing out every "maybe" became quickly obvious to early users of such machines; faced with the examination of mountains of paper output, they soon gave orders to the program to "tell me only if such and such a threshold of expectancy is reached".

An even more sophisticated arrangement could be made for Test B to be run whenever Test A produced a certain result.

Any success in "explaining" the phenomenon may depend on the experience and savvy of someone more senior than the "finder". Every practicing cryppie has a lot of "effect and cause" data tucked away in a mental notebook, and is likely to recognize an old friend (or distant enemy) by little behavior patterns, much as a detective builds up "profiles" and sometimes recognizes the handiwork of a previously encountered criminal.

If it is a brand-new case, the analyst (just as the sleuth) is expected to come up with at least a possible explanation. Exploitation of the phenomenon implies an actual solution, either partial or total.

In "real life" one may have to settle for something which nudges a difficult problem at least one small step forward.

In our more recreational enterprises, it is usually an all-or-nothing proposition, with a complete answer to whatever question is posed, such as the final plain-text version of an encipherment.

Without dwelling on the matter, it is reasonable to assume that (again) in "real life" the fourth "exploitation" phase includes somebody someplace doing whatever is required to put to good use whatever information has been supplied by the cryppie — who may not know (or even care) what happens at this next step.

The exploitation naturally depends on the degree of difficulty of solution, the volume of traffic, and a host of other considerations — including the resources of the person or agency concerned with the problem.

It may boil down to a question of "is it worth it", or "do we want this", so a wide area of outside interests may have some bearing on the matter at hand.

The point may be of some interest that it is not necessarily easier to "keep current" (as far as exploitation is concerned) with basically "simpler" problems than with what should be considered "more difficult" systems.

This is due to the fact that "simpler" problems may have to be tackled on an individual basis; you can almost guarantee a solution will be made sooner or later, but you cannot necessarily guarantee that it will happen today or tomorrow or next week.

On the other hand, even though it may have taken months or even years to have solved the mysteries of wiring, stepping, daily keys, setting keys and so forth of a complicated cipher machine, it is possible that once you have licked the beast you may

Perhaps daily keys will have to recovered (in the wee hours of the morning?) or a recovery on some regular basis will have to be made to determine which set of what components (previously recovered) are currently being used.

In a way, it is too bad that there is no "puzzle" type of problem in the public realm that offers the kind of challenge that solution of indicators and setting keys provide.

For it is obvious that without successful mastery and exploitation of these seemingly "external" security devices, the solution of a specific machine or a complicated encoding procedure might still produce but a trickle of information — most of which might be old enough to be useless.

It may be that feats like solving next years anticipated keys (by discovering a method of creation and a rule of subsequent use) are the ones deserving of medals to match those bestowed on the system-or-machine solver.

But in a chapter on "diagnostics," it seems perfectly reasonable to present phenomena that relate to key or "rules of behavior" in addition to the "funny things" in the cipher itself.

To sum up, in some tough cases you may be either all the way in or all the way out — either batting 1.000 or setting strike-out records. And at the same time some relatively minor systems will chug along on a very erratic course — with a pretty constant record of 10 percent solution within so many hours, 50 percent within a day, and so on.

Oftentimes it will depend on the use of certain components for a certain length of assigned usage, wherein each month (or whatever) the solution rate goes from zero to a hundred percent, but hits the bottom of the cycle at periodic intervals while the poor cryppies are struggling to reach the top of the well for the umpteenth time.

But let's get back to diagnostics.

Finding the Phenomena

In order to discuss the finding of "something funny", we had better try to define the sort of thing we are searching for. The following is far from a complete categorization of typical cryptographic phenomena, but represents a loose structure on which we can pin certain labels, and say that such and such a "thing" which we may have noticed fits within certain sub-divisions, perhaps in multiple associations. By doing this, we then may be able to "fingerprint" a case, and match it up with previously understood cases.

If we can "label" the unknown properties we can also "label" the properties we know to be present in recognized systems.

I am going to use that dirty word "OBVIOUS" for one category — but I only mean it in the sense that certain phenomena just sit there staring at you. You don't have to

I am going to use that dirty word "OBVIOUS" for one category — but I only mean it in the sense that certain phenomena just sit there staring at you. You don't have to perform any operation (except to examine your material) in order to make things apparent.

As opposed to this, there are certain properties which will only become apparent if you "do something", such as make a comparison between two positions, or look at every third letter, or perform some arithmetical operation on some cipher element. We will call this second category "LATENT".

Within the OBVIOUS and LATENT categories, we may observe certain things to which we will also apply names. Sometimes it will be a specific and recognizable "something", such as the date of the message, or a check on the group count, or a message classification.

For lack of a better term, we will call any such recognizable word, or group, or number an IDENTITY — something that stands on its own, with a meaning or function you can talk about. You might see this in a negative form, in that there will be no letter J in a sample large enough to contain several, according to an otherwise fairly random distribution.

But many times what you notice will not be an IDENTITY (standing on its own) but a RELATIONSHIP between two elements of the cipher.

This may be a presence or absence of repeats — adjacent or at some fixed interval.

If there is some repetitive process at a fixed interval, we may use the term CYCLE to indicate a relationship which involves a repetitive key or something similar.

If you are lucky enough to have a crib, there may be an obvious RELATIONSHIP between cipher and plain, in that a plain value will never be found associated with the same cipher value — E to E, for example.

A very common thing you will be looking for is DISTRIBUTION, which may be phenomenally rough or phenomenally flat — or comfortably within the random range. This can be overall, or taken at a fixed interval, or cycle.

A very important point: Relationships may often be found between messages, rather than within messages. They may be obvious hits, or they may be latent — for instance consistent and repeating differences between certain groups of message A with message B, when they are properly superimposed.

Or perhaps there will be a relationship between groups 1 and 2 of every message which show some property when properly examined — either a constant, (such as a common address or security classification) or a progression (such as a message number).

Sometimes it is not so much a case of seeing "something funny" as it is of seeing a very specific funniness — tantamount to your not being surprised that in message after message you always had a few letters at the low end of your frequency count, but your interest quickened when you realized that the letter E was one which was always low.

If it had been D, or P, or even Q, you probably wouldn't have reacted. But why E — the old friend we so frequently look for? You are entitled to add some plus factor into whatever statistical weight you gave the absence, although it is difficult to put this in scientific mathematical terms.

For isn't it true that a reflecting type of device, such as the ENIGMA, wherein a plain letter cannot be that cipher value (under usual conditions) would perhaps have a slightly biased count, if you had enough traffic? Naturally the above story wouldn't be very exciting to a statistician, since the effect may be very, very slight.

Explaining the Phenomena

Let's take a more striking case: Pretend that although we do not have enough traffic to make the distribution absolutely "phenomenal," there seems to be a trace of roughness on the first few positions of each message when lined up.

While most of the frequency counts have a reasonable curve, from a slightly higher-than-expected occurrence to a few completely absent letters, one could consider such "blanks" as being within the boundaries of possible events.

However, as we said, sometimes it is not so much the fact that "something" happens; it may be that our interest is in "exactly what happens."

Let's say that the system has all the trappings of Playfair — no J, (or a combined count of I and J which roughly matches the rest of the letters), and no repeats within digraphs (AA, BB etc.) with the expected amount of repeats across digraphs (DA, AR, ET TX, etc.)

But out of the very few repeats you have between the beginnings of messages, on one day you see three cases of a beginning sixletter repeat:

1. **CRSEB OLCFB** .....
2. **CRSEB OYTTX** .....
3. **CRSEB OPAFG** .....

Can you say anything to elaborate on a statement that IF this is Playfair, and IF such and such happens with the first two letters, then THIS event MUST happen? If so, it might give quite a twist to the not-very-impressive odds of your finding a triplet of 6-letter repeats at the beginnings of messages.

Now I have to confess that I left this "duplicate problem" in, merely to give you the opportunity of checking your absorption rate. Do you recognize a twist on the tale about finding a pair of messages starting out MBNU — in the NOTES to the last Chapter?

You probably are a good student, and caught both the duplication in my story-telling and the "secret" behind it. But occasional reiteration of an important principal may be justified.

Of course the fact that the first digraph of a message hits with another message is only a one out of 676 matter — or 525 with no J present, and squaring that number for the purpose of figuring the probability of a four-letter beginning hit may not be enough to get one excited if we have oodles of traffic.

But the thing is, this may be an interestingly specific digraph to hit, — or the SE of the second digraph might be considered as also being interesting. More precisely, what if this IS Playfair, and just by chance (one in 525 since we have a 25-letter alphabet) the square was so constructed that SE was enciphered by CR?

(Now comes the kicker.) What MUST happen in this system when you encipher CR (unless those letters are in the same row or column)?

The answer is obviously SE. And because it HAD to happen, and we DID find it, the cold statistics of one-out-of-525 digraphic repeats are temporarily tossed out the window. Again, IF it is Playfair this WILL happen, which doesn't by itself prove it IS Playfair, but gives us a terrific boost on the way toward a very serious attack.

So what we saw was an OBVIOUS IDENTITY — but it needed a catalytic agent in the form of an analyst familiar with the properties of Playfair before it became truly phenomenal.

Now let us see how our nomenclature would fit a few typical systems we have talked about. No attempt is made to completely exhaust the possibilities—we may just consider a few things likely to occur. A more complete tabulation might show many "odd" things to expect.

Playfair: You should expect (in the OBVIOUS category) no J. Repeats will probably occur, on an obvious cycle of 2 — and the digraphic distribution will probably be rough.

In this distribution there will be no doublets AA, BB, etc. In case of a crib, matching letters between cipher and plain will be zero.

Quagmire: Some OBVIOUS phenomena, such as repeats, an IC a bit above random possibly, but many LATENT characteristics. These will include the establishment of a cycle by either rough distribution or repeats at a specific interval.

With a crib, relationship between cipher and plain will be consistent on a cycle.

Transposition: Most phenomena OBVIOUS — cipher will have plain text distribution, but there will be a negative IDENTITY of repeats in all likelihood.

Bifid: Apart from the probable absence of J (as occurs frequently in a 25-letter alphabet) most phenomena will be LATENT. There may be broken repeats which will be made significant at the correct cycle.

With cribs, there will be consistent relationship in each of two cuts between plain and cipher values at least to a definable extent: Although plain E will not necessarily be over the same cipher value—row or column — even when written on the correct cycle, EN, for example, must have a constant relationship for either a row or column value.

Overall distribution may be somewhat rough, with possibly a different roughness for the "top" and "bottom" streams of a writeout. (This is due to the fact that we are seeing nothing but plain row values on top, and plain column values on the bottom level.)

Autokey: Probably no phenomenon of obvious distribution, but a latent phenomenon of a biased Delta count, which is the difference between consecutive cipher letters.

If a mixed alphabet is employed this effect will be harder to establish, but with known components one can achieve the same effect as with normal alphabets. In the case of plain text autokey and a crib, the length of repeat may be one letter short in the cipher — THE, for example, may turn up as a repeating digraph. In plain-text Autokey, the fact that the key is rough (since it has plain frequency distribution) may leave a trace of statistical roughness in cipher.

Now let us consider a few unknown cases — which may or may not be "cipher systems." (I have included several little problems of spotting what is "funny" wherein you might try and figure out what could have caused the phenomena you are viewing.)

Case 1: 29046 12546 29678 30607 88610 32947 45272

How would you describe any phenomenon you see?

Case 2:Message: **TCHAC HTACH ATWBI AWKEA BWILI KIBKW**
CIHBW BEBWK TLCEA THCIL IKIBW KTAEC
TAWKI LTCHA HCATC etc.

How do you explain what you see?

```
Case 3:Message:  RENSR SUSPC NSESE ESURS RSRES EBNSR SLSUC
                 ISRSR USRSR SRSEN BSRSR RSRUI SESRN UESRS
                 RLSAS ESRSE ASIAP SRSEI SRENS RSUSL etc.
```

What can you tell about the system?

```
Case 4: (With crib)          P L E O F F I
              Cipher: H F A S R M M U N I R F S O C F K N
```

What do you think?

Case 5: The first three and last three groups of a brand new system:

```
87574 97577 28901.....36104 86085 98688
19684 30660 25578.....86493 80021 64811
69539 12053 43487.....44647 83754 22262
35908 24768 21993.....82328 62196 21139
10824 77835 62743.....55991 31964 65415
```

How do you know that you probably have the first five messages sent?

If you have had time to make a cursory inspection of the above examples, you will probably have spotted most of the phenomena, and may have a logical explanation for the cause.

In case there was any slight difficulty, let's examine the cases in turn.

```
Case 1: 29046 12546 29678 30607 88610 32947 45272
```

You may not have needed to explore this so specifically, but let's take a frequency count — writing the second digit of each running pair under the first digit:

```
0 1 2 3 4 5 6 7 8 9
4 2 9 0 6 4 1 8 3 0
6 0 5 2 6 2 2 8 8 6
7   9   7   7 4 6 4
3   9   5   0 2
    7       1
    x
```

What is the order of frequency? (2 6) (0 4 7) (8 9) (1 3 5) There seems to be an obvious bias on the even side. Actually there are 22 even and 13 odd numbers.

Just for the fun of it, let's take a count of Odd-Even, Even-Odd, etc.

Odd-Odd	Odd-Even	Even-Odd	Even-Even
None.	13.	13	8

This distribution is much more phenomenal than the one above, peculiar as the even-odd split might be. Given an odd digit you MUST have a following even digit.

(The even digits are followed about randomly even or odd, given the biased overall count.)

Of course you probably said this long ago, when you noticed that there was a rhythm to the digraphic count — 29 04 61 25, etc. In other words, all the digraphs begin with an even digit. Perhaps it implies a 5 x 10 matrix with room for common values, punctuation, etc.? Or it could even be a simple 5 x 5 square, with only the even digits as row names, but two alternatives for column names.

(This could have been a rather tricky problem had a 5 x 5 square been used with five randomly selected odd/or/even rows and all ten digits used for columns. You would spot it on enough material, but the short sample given might not have been enough.)

One minor suggestion: It is just as easy to divide your count by taking a tally on the cut of two, and at the same time over the cut as a separate tabulation. That way, you can see if there is any on-the-cut property (as in Playfair) and still combine the counts if they show no distinction.

They may either match, or else give you a clue as to what to do as the next step.

Let's examine Case 2. Obviously this count is severely limited. We could take the time to make a specific frequency, but since there are only ten letters used, that in itself might be a preliminary fact to digest. Why 10? Suggests numbers, perhaps?

Lets take a running count, by writing down the following letter under the first member of any pair. But, just to save time later, perhaps, let's break the count into two sections, 1-2 as part of the first distribution, 2-3 as part of the second, then repeat the process — 3-4 as first distribution, 4-5 as second, and so on.

As we pointed out (deliberately too late) in the first example, the slight "trick" in doing this is that just in case the system reveals itself as being digraphic, you will have used your pencil efficiently.

Alternatively, you could make a partial count of 1-2, 3-4, 5-6 etc., and look at it, ready to make a second count if it looks like a good idea. For one thing, you would want to know if the two counts are similar.)

Let's try that 1-2, 1-2 routine first.

Eager Beavers who glance at the tabulation of the ten letters present (and say "what's the next problem") should be patient while we proceed slowly and carefully for the benefit of the less experienced.

T	H	C	A	W	I	E	B	K
C	A	H	T	B	A	A	W	W
A	B	H	T	K	L	B		
L	C	I	T	B	K	C		
A	A	E		K	B			
A	C			K	L			
C				K	K			
					B			
					L			

So that's it. OBVIOUS DISTRIBUTION. There are two classes of letters — A L K B C and H I E W T, and a member of one class is invariably paired with a member of the other. However, the digraphs seem to be reversible, in that you can start with either class in the first position. (Whether TA is the same as AT remains to be seen , but one might expect that it probably is.)

With two such obvious words to fit as coordinates of a 5 x 5 square, you would have no problem solving this as a monoalphabet. Actually, if you "tried the simplest thing first" you would hardly call it "solving."

The only slight catch is that you might have to try BLACK as either a row or a column designation, unless you looked closely at the frequency count, and have decided that reading it in one very specific way you got a Normal alphabetic frequency count, from A to Z (no J of course).

A hint might be that at one blind trial you get the message beginning as TMR, while the other way you get THE — so there isn't much of a problem in choice.

Of course, in such a problem the square can be mixed in many ways, making for a less "phenomenal" and less easy-to-solve example.

Now for Case 3. As in the above, the count is very rough, but this time there seems to be 11 letters. From experience, I would suspect numbers and a separator of some kind, but in any event, you might want to take a frequency count of some kind.

It won't do any harm to take a digraphic count, even though you could jump into a more sophisticated test if you happened to think of a reason for the 11 values. Incidentally, there is a "glitch" in the encipherment — or a garble, or whatever. But I left it in to show you that it is possible to arrive at a sound conclusion even with shoddy material. Remember the makers of "problems" in the magazines sometimes make a mistake — and we all know what radio transmission might do to a spy message.

Again, experienced anagrammers, please bide your time.

```
R E N S U P C B L I A
E E S U S C I N
S E S E S
S E   R
E B   L
S N   U
S     R
S     R
```

Just indexing less than half of the message, on the 1-2, 3-4 etc. stream, shows heavy bias for certain letters. Just to check, take a little count on the rest of the message on the "other" beat (2-3, 4-5, etc.) — although I started around the group SRSEN.

```
R E N S     U P C B L I A
S B R ERI   E       S S S
U S   RIO             S P
N S   ERR
S A   REU
  N   RRL (To save space, I added the S count
      R    as a "clump" - there are so many.)
```

Same general shape, but even heavier on S. Now the perspicacious types who have been kicking me under the table, saying they know the Keyword, should relax while we go through the motions of proving that S is a separator. Rewrite the message, leaving out S, and see what it looks like.

```
REN R U PCN E EE UR R RE EBN R L UCI R RU R R R ENB

R RR RUI E RNUE R RL A E R EA IAP R EI REN R U L etc.
```

I have numbered the groupings in order to point out something — an experienced analyst would see it "in his head." As a matter of fact, the enciphering clerk obviously made a mistake at one point, but I want you to spot that, as well.

```
 1   2 3   4   5   6 7   8 9   10   11 12 13 14 15 16 17 18
REN R U PCN E EE UR R RE EBN R  L  UCI R RU R  R  R

19  20 21 22  23 24  25 26 27 28 29 30 31  32 33 34  35
ENB R  RR RUI E RNUE R  RL A  E  R  EA IAP R  EI REN R U L
etc.
```

Notice we have mainly clusters of 1, 2 or 3 "digits" (if that's what they are) plus one maverick. When we examine them closely, we notice that the three-wide groups are in positions 1, 4, 10, 13, 19, 22, 31 and 34 — and the maverick is in 24. (The counts past the maverick perhaps should be considered suspect, for the time being.)

If you are used to cycles, or are a math student, or just on the ball, you have probably spotted the fact that there is a rhythm to these 3-wide groupings — they always start

a cycle of 3. So let us write the stream on that basis, forgetting the position number this time.

```
REN R U   PCN E EE   UR R RE   EBN R L   UCI R RU   R R R

ENB R RR  RUI E (RNUE)  R RL A   E R EA   IAP R EI   REN R U
```

Now the interesting thing becomes the middle letter of the trio of groupings. It is always R or E (except for the area right around the suspect 4-letter combination. I suppose we are ready for a hypothesis, and perhaps even a peek at the obvious keyword.

Let's speculate for a moment on what we would have said in an entirely different situation, where the groupings were in twos instead of threes, and the number of letters in the first member of a group-of-two-clusters could be 1, 2 or 3 letters, but the final member was limited to two (digits) with a severe limitation on the first member of any cluster. An experienced analyst would have probably quickly said "book or dictionary code, giving page and position-on-page".

Thus, THE might be the 38th word on page 572, so he would have sent the correct letters for 572 followed by 38. "A" might be 81 if there were 7 pages of preface material.

Obviously a page in a reasonably sized book can go as high as 700 or so — unless it is a large tome, there may be no 900 pages, or even 800. But if you are marking off lines on a page, for example, or entries in a dictionary, the chances are that you will be limited to 30 or 40 on each page.

The limitations on the first digits will quickly show you which letters stand for the higher digits 7, 8 and 9, and the line numbers will probably point out which letters are 1 and 2, even if you can't anagram anything.

But we have three members to a set. Most small dictionaries have two columns on each page. Rather than laboriously count up to 38 or 40 or however many entries there may be on a page, it might be quicker for an enciphering clerk to tell you whether he was using the first or second column — and only have to count half as high to get his proper designation of a word.

That would explain why we get nothing but R and E as the middle member of a grouping of three "things". First or second column on the page.

So we know our keyword starts RE (if the numbering is 1 to 0) or possibly -RE (if 0 to 9). Now you eagle-eyed cryppies who have been hollering REPUBLICAN for the last several minutes can ease off.

Putting the message in numerical form using:

```
          R E P U B L I C A N (S is separator)
          1 2 3 4 5 6 7 8 9 0

we get:     120 1 4   380 2 22    41 1 12   250 1 6 etc.
```

You wouldn't be expect to solve this little scrap as a "code" problem, but if you had enough material you could either "break" the code or guess the book.

(We remember many long hours spent searching the shelves of the Library of Congress for out-of-print dictionaries in some foreign language.)

In a real case you might have to decide whether the method is Page-Line-Position on Line (for either a word or a letter), or is something such as we suggested.

How about Case 4? We are lucky enough to have a crib:

```
                        P L E O F F I
Cipher:      H F A S R M M U N I R F S O C F K N
```

Those who are used to juggling the numerical values of letters would see (in effect):

```
                    16 12  5 15  6  6  9
 8  6  1 19 18 13 13 21 14  9 18  6 19 15  3 6 11 14
```

Notice anything slightly unusual? Not absolutely phenomenal, but something in the order of tossing ten heads or so before you got a tail.

In any event, if you felt the difference between the cipher and plain values looked limited, it would be worth writing them down:

```
2 1 8 6 8 3 9
```

Notice anything else? This is one of those cases that MIGHT be important, and a simple test would either bear out a hunch, or send you back to the drawing board. We talked earlier about Fibonacci, and the series generated to be used as key. We see that on a cycle of 5 (a nice odd number.) 2 and 1 make 3, 1 and 8 make 9. Dare we try the next few numbers we would generate, such as 4 4 1 2?

If we are so bold, we get B O N A — extending out plain to P L E O F F I B O N A — which should certainly encourage you to consider this an "example of Fibonacci". Incidentally, it ain't that easy to work backwards if you have only 5 digits of key, so you would have to go forward to test, before getting enough key to "back up." It's much nicer to have a crib confirmed.

Case 5 represents not a message to solve, but a very important part of the cryptanalyst's work. The phenomenon involves the first and last groups, or rather part of them

```
87574 97577 28901.....36104 86085 98688
19684 30660 25578.....86493 80021 64811
69539 12053 43487.....44647 83754 22262
35908 24768 21993.....82328 62196 21139
10824 77835 62743.....55991 31964 65415
```

The easiest way to see whether or not something remarkable is going on, may be to let your eye wander a bit, scanning across rows and down columns. You may stop for

a moment or so whenever things look interesting enough to warrant a real investigation. (Have your pencil ready, but don't immediately start tallying blindly.)

Look for unusual presence of or absence of digits repeated next to each other, or repeated trinomes, etc.

You stop for a second at the 757 repeat, see no similar pattern in other groups or the same columns of other messages, and move on. You know by now that seeing such strings as the 8901 consecutive run is probably a chance occurrence, since nothing unusual happens in that vein.

We have something less than 150 chances to make comparisons of consecutive digits, and it is interesting to note that a quick count of repeats of neighbors shows 14 — where we expect 15 as the ideal chance number. Not much excitement there.

I like to take several small samples and combine them, when taking a frequency count — just in case there is a local aberration. Doing this with the above, my counts were:

0	1	2	3	4	5	6	7	8	9
7	6	3	3	5	5	8	5	12	5

0	1	2	3	4	5	6	7	8	8
2	5	11	8	6	4	6	4	6	7

0	1	2	3	4	5	6	7	8	9
1	4	2	3	4	5	3	3	2	3

They may look rough in small sample form, but when we put them together our total is:

0	1	2	3	4	5	6	7	8	9
10	15	16	14	15	14	17	12	20	15

I missed a couple of digits someplace, but it isn't important enough for a recount. Proving the distribution is pretty darn flat, I took a Phi count — n(n-1) + n(n-1)..... + n(n-1) divided by N * N-1 where N is the total frequency —and I get 2097 over 21756.

If that doesn't work out pretty close to the 10% random figure, you can say I should hire a mathematician.

So now I've led you up a few garden paths to show you what you MIGHT do. But if your eye dropped DOWN the columns in various places, as well as across the rows, you might have been intrigued by the fact that the sequence down column 1 — 81631, is repeated down the penultimate column.

When you try other places, the only exciting fact is that the second column has a relationship to the last. Put another way, the first two and last two positions of each group have a very interesting phenomenon:

87	88
19	11
69	62
35	39
10	15

The differences (non-carrying, in the customary manner) are 01 02 03 04 and 05. Can you think of a more simple disguise for the message number?

You won't get very far trying to "read" this short sample of text, so don't waste your time. And you may be perplexed by the inclusion of such a seemingly irrelevant problem in this diagnostics chapter. But suppose the phenomenon had been one of disguising the indicator, or setting within a book additive system, or something like that.

The exercise was to point out that relationships between the first and last groups, or first and second, (or in fact any two — or four — groups) may yield a gold mine of information if you are faced with setting messages at different places of a key, or are looking for settings of rotors on a cipher machine, or need a discriminant to distinguish one kind of traffic from another.

A final word as to how one might go about looking for phenomena in an orderly way — perhaps using electronic assistance.

Almost all of that sizable segment of not-necessarily-professional solvers who use computers have their own pet programs for either general cryptanalytic procedures, including diagnosis, or for solving specific systems — after such use has been determined.

I and some others term the all-embracing program "Interact" — and the object is to include many options that you can call for upon demand. Ideally, your original diagnostic program will suggest a certain category of system, and you can then press a specific key and the subroutine will take over.

Of course, most of the problems ACA members are given are labeled QUAGMIRE II, or Playfair, or whatever, so the initial steps of looking for unknown phenomena are not necessary. The most frequently used tests are therefore geared to such things as looking for repeats or cycles, and taking specific frequency counts (such as those on a specific position of a specific cycle).

But it may be helpful to examine what is built into the small battery of diagnostic-cum-exploitive tests such as I have for my reasonably powerful home computer. (The big "computer nuts" in the organization have much more fancy machines and programs.)

As one starts to run the program, the first thing the computer asks is whether you want to examine some previously saved text, or are you starting a new job. The ability to "save" a message input (perhaps as soon as you have entered the data) is important, since quite often you do not make an immediate conquest, and might want to come back to a toughie at some other time.

The next question the computer asks, in effect, is "what language are we thinking in?" The Order of Frequency of most of the major languages is stored, including such as Interlingua and Esperanto.

Although they are not part of the primary Interact program, I have samples of various text stored away in a "Help" (Language) routine, which can be of intermittent assistance only, since I do not have the capability of looking at two things (or windows) at the same moment — or at least I don't have the program to effect this.

Having answered the question concerning the desired language, I am now instructed to enter whatever message text I desire, with or without word separation. At present, I am limited to 280 letters at a time, but this could be changed if I want to sacrifice the ability to look at a lot of things at once, such as the results of certain tests, the names of the many options available (and how to bring them in) etc.

After I have punched in text, I often press the designated key (D for Disc save)to store the text. (The computer asks for an identifying name or number.)

The "beast" now waits for me to order the next job, which usually means entering M (for Multiloop) upon which the screen immediately shows all the 5-letter repeats and the positions at which they occur, followed by the 4-letter and the 3-letter hits. At present I have to do the subtraction (distance between hits) mentally, but this could easily be written in.

But hits are only part of the test for cycles, so when I enter S (for Stethoscope) the machine starts to take a count at all the various intervals (presently up to ten — although this could be expanded). After a short interval of activity, the screen tells me that I should consider such-and-such a width, giving me the average I.C. at that interval. Of course it merely shows me the "best" one, which may not be any significant figure if there is no true interval.

If I so desire, I can now enter P (for polyalphabet) and the screen will show the text written out 9-wide, or whatever, with a space between groupings. Each of the "columns" (on a 9-wide writeout) is a separate entity, and can be manipulated independently of the others.

The screen is rather full, since it shows almost all of the the important things as a basic picture — the Plain and Cipher alphabets as they currently stand, the cipher text (at correct width if you have determined such), a box for the frequency of letters from A to Z at whatever stage you happen to be in at the moment, the I.C., and a table of twenty or so options (such as (S)tethoscope, (D)isc save, (M)ultiloop, etc.)

There is also an area for intercommunication in which it prints (Alphabet?) and I enter the 1, or whatever figure I want. When Multiloop are some other subroutine has to display something, it waits in that position until I press the Return to the standard display.

To explain about "solution" procedures, it may be simpler to go back to the point where I ordered the polyalphabetic test. Suppose the message is obviously monoalphabetic. The entry may or may not have been with word division.

By entering T (for Try) the program will try to solve the monoalphabet by assigning E to the highest frequency letter, T the next, and so on. Naturally with less than 280 letters, such a "brute force" technique is unlikely to produce much "plain text," but it

may be surprising to see that the first three letters come out to be TDE, or something like that.

The putative plain is entered immediately under the cipher stream. Also whatever substitution is performed shows in the Cipher alphabet, which appears at the top of the screen under a Plain normal alphabet.)

Assuming such an instance, to "correct" the plain, all I have to do is enter X (for eXchange). The screen then asks me to enter the "wrong" and the "right" letters, and immediately all the Ds in the plain version become H, and all the Hs become D.

This exchange can be limited to specific alphabets when you are dealing with Quagmire types. For example, you could change all the Ts to E (and vice versa) in the third alphabet without disturbing the other columns.

There are several other program controls and housekeeping functions built in, such as the obvious ones for Q(uit), (C)ipher Edit (if you want to add or correct) and the possibilities for further sub-programs are only limited by the capacity of the machine and the ability of the programmer.

A digraphic program, modification for digital text, and other routines can be worked into Interact, and there is a constant stream of programs applying to specific systems coming from the ACA computer group. (In addition to the Computer Column in the regular issues of *The Cryptogram*, the "computer group" has its own — and independently paid for — Computer Supplement of many pages.)

For those without electronic aids, I trust that the steps outlined might aid the pencil pusher in devising an orderly procedure for attacking known or unknown cipher problems.

Now that we have tackled many known cipher types, and seen how to analyze the unknown, I trust that the reader will have some fun with the "Problems to Solve" given in Chapter XIV, after a last look at a few maverick systems.

NOTES ON CHAPTER XII

(1) You know by now that "Some Of My Best Friends Are Mathematicians" (the title of a spiel I have often given to the Learned Brethren) and we have a few good-natured tiffs now and then.

I sometimes accuse them of trying so hard to fit in some pet formula that they may be overlooking a very obvious "something" that to a layman would present less of an obstacle.

I have had occasion to give the same problem to a number of mathematicians and have sometimes been surprised at the wide variance of answers, and the span of time elapsing between that of a sharpie who immediately gave me a satisfactory estimate (which may have been all I needed) and the several-days-later hand-in of a very careful worrier — many sheets of closely packed figures.

Perhaps one of the reasons for my surprise lies in the fact that there seem to be several families within the genera within the species of mathematician.

If I query one promising professional, he or she may respond "Oh, I'm a geometrician — you want an algebraist." And I seem to stumble upon a topologist when I have a problem in statistics, or a "pure" type when I want a garden variety of "count 'em and tell me" answer.

And my particular bane is the occasional union-card-carrier who expects to be brought all the facts and figures for consultative consideration, in specified arrangement, whereupon such a "professional" mathematician will deign to give a considered opinion.

In our business everyone was expected to root through raw data, push whatever pencils were necessary in getting things organized, and be capable of handling an entire job from start to finish.

Having said all that, I will tip my hat to the many outstanding members of the order who have done so much in areas where a layman simply cannot dare to tread. And most of them have been hard workers, willing to do anything without thought of who does the slave labor and who gets the credit.

But I started to tell a story to illustrate my tenet that sometimes it pays to sit back and visualize the problem as a whole, looking for "ways to pick it up", and perhaps venturing a "common sense" approach rather than automatically looking for a formula, or a "general solution" to a specific question.

There may be times when a person who professes no particularly mathematical skills may immediately get to the heart of what seems to be a "mathematical problem" at the same time a much more erudite "pure" type is trying to fit so-and-so's theory to this particular case, or adapting a complicated formula to the few figures available.

One of my sons was a boy soprano in the very good St. Alban's choir in Washington. At an early age he had displayed an aptitude for mathematics, and found it a convenient area for exploration during the less exciting moments of a long service. He might sit and think about various problems, and dream up a question to ask the extremely competent professionals he sometimes met in our carpool, or when they dropped in after hours.

One of the problems he posed consists of determining the number of oranges that can be packed around a single orange so that they each touch the center piece — all of the outside "touchers" being the same size as the inner and "touched" orange. (Use tennis balls if you'd rather.)

I gave the recreational puzzle to several mathematicians for lunch-time consideration, and when I began to get solutions from them, I was taken aback by the wide range of answers. Remember, the guys and gals I asked were by professional standards considered rather proficient at their trade.

Naturally no one worked hard at getting an answer, and some of the "guesses" were probably rather wild shots in the dark. But when I tabulated answers from a low of 6 to a high of 36, or thereabouts, I began to think about methodology.

If you want to consider the problem yourself, think about it for a while before you read on. Naturally a few came up with the mathematically correct answer of a number someplace between 12 and 13. These solvers probably used a fancy formula, and knew how to use it.

But when I checked my hunch by asking a few non-mathematicians for the answer, they seemed to be scattered about in the same way as the "professional" returns — some rather wild extremes and a few on the mark. Perhaps this proves that I wasn't asking a mathematical question. Now to enlarge on my speech about "how to pick it up." Certain people seem to be able to visualize this problem, and actually see a ball hanging there, with a cluster of hangers-on. Maybe all they have to do is count what they see.

I happen to think the easiest way to solve this quiz is to think about what we are trying to count, and see if there is some analogy in our experience. Doing this, I thought of two ways to arrive at an answer — one of which sounds reasonable, but might be hard to prove.

The best way (for me) ended up by visualizing part of the problem as a plane. When you were a kid, didn't you play around with a compass — I mean the type that swung around on one fixed point on a piece of paper, with a stub of a pencil stuck in the freeswinging "leg." (I probably should say "pair of compasses" to avoid any confusion with the Boy Scout's friend.)

On a plane surface, how many circles can you make in your doodling with the compasses, which all touch the outside of the circle (made with the same compasses setting?) More aptly, perhaps, haven't you some vague idea about radii, and pi, and circumference, and stuff like that, which seem to ring a bell with a number six attached?

Any way, I think it is fairly easy to visualize a ring of tennis balls that fit around our object ball, going around the equator, so to speak. If our compasses rules have not been changed, there should be six such spheres in the ring — all touching the middle ball, whether or not we care about their touching each other.

Now think about any "meridian" of the orb we see hanging there. Obviously the "equator ring of six" crosses this meridian belt (which runs around both poles) twice, and so we have already placed two balls on the meridian.

Doesn't that mean we can only add four balls for our first meridian attempt? We have thus added four to the stock of six on the equator, making ten packed so far. But isn't there room for more?

Let us try to visualize another meridian to fit in someplace. Obviously we can't put anything new on the equator — and equally obvious is the fact that we can't put anything at the poles. (Our first meridian took

those parking spots.) But there seems to be enough room for two more, to form a new meridian which passes through the so-far uncrossed balls on the equator and the already placed poles.

So six, plus four, plus two equals 12 — and there ain't no room left. Well, there is a little loose space, but not enough for another ball.

The second way to think about the problem as a visual exercise consists of imagining the two Tropics on the earth-like globe we see suspended.

Obviously if we can put six around the equator, it's too much of a stretch to put six around a belt much further removed — but maybe we can put in five close to the Tropic of Cancer, so to speak, and then hope to zigzag an arrangement of five more below the imagined Tropic of Capricorn.

It seems to be possible, but if we put a ball at each pole, there doesn't seem to be enough room left for any more. So we end up with five and five and two — the magic number of 12 again.

Now I don't pretend that this is mathematics, or even demonstrable apart from getting out 13 tennis balls and packing 12 of them in the right-size box, with sawdust or something to hold things in place. (And we can't find room for the 13th.)

But it illustrates how some people, who wouldn't recognize a Sigma sign if they saw one, can come up with a "kind-of" solution which someone more scientific could use as a basis for exact reasoning.

In the same way, some people "visualize" how a cipher machine transfers its current through a maze of rotors, and comes out with a specific cipher letter for any given plain entry. Such a person mutters to himself "It comes in here and goes out here" — and the music goes 'round.

Or he or she will write down the equivalent of a rotor as a "box" or "tableau" and trace the exits and entrances just as they would a Quagmire encipherment.

Engineers, on the other hand think in terms of wiring diagrams and hot points and such, while a mathematician is likely to look for some formula full of X's and Y's which can be used to illustrate any given point of encipherment.

I will never forget an incident that happened while I was giving a course on a commercial enciphering device to a group of "special students" — some of whom were outstanding analysts, or mathematicians, or whatever, but were just getting a little exposure into areas they may have not dealt with before.

As is usual in such cases, after covering the encipherment process rather thoroughly, I started the "game of exercises" by expecting them to recover certain plain text when they had the advantage of knowing most of the components — such as the daily setting, or the position of certain integral parts, etc.

Then we progressively went through several more quizzes, where less and less was given and more and more had to be recovered. Finally we got to the point where a general solution is required — usually a fairly complicated procedure compared with the first situation where almost everything is known.

One of the "students" in the class (if I can call them that) was a recognizedly brilliant mathematician, who later was given accolades for certain work but for now was just another naval officer. I looked at his worksheets particularly, since he had seemed to take longer on the first "obvious" and easy cases, but finished the "hard" part of the exercises much faster than anyone else.

Lo and behold, he was using the same method for all the problems — a general solution, even though it ignored much of the "given material" in the easy cases. To him, it was much the simpler job to apply whatever formulae he threw at the case than it would have been to constantly adjust to "crutches" and say "all I have to do is think of the electricity going in here and coming out at this point."

I'm sure my mathematical-genius-student would not understand my amazement, and might be surprised to learn that I regard his "general solution" with so much fear and trepidation.

(2) As you know, I can't be very specific about details of any successes we may have had during World War II. But enough has been written (or implied) to indicate that there must have been some such.

The exploitation phase of cryptanalysis was far removed from the area where pencil-pushers were scratching their heads and scribbling on cross-section paper. Naturally we had some inkling of the use to which our efforts were being put, and occasionally got secret delight out of seeing a newsreel at the local movie portraying a successful landing at the part of some island where the enemy wasn't (instead of our forces stumbling onto a hornet's nest at the part that they were). And the shooting down of Yamamoto (with both Army and Navy reading of his intended route) carried special significance beyond the implied fact that a couple of fighter planes happened to be in the area.

(3) Without naming names, I would like to underline the fact that a team enterprise demands the best being given on all levels, including those who only do the pick-and-shovel work of cryptanalysis — of which there is an ample quantity.

One of our first major breaks into a new and difficult system came about when it did, because an enlisted man in the far field thought he "saw something funny" and doggedly pursued a routine of checking and cross-checking until he was convinced that an unexplainable relationship existed between certain digits of certain groups.

I emphasize "unexplainable" — at least to the lowly clerk, whose experience would hardly place him as a candidate for their best diagnostic team. But he did have enough perspicacity to politely halt the first officer who came down the aisle, and show him his findings.

It may be that the young Lieutenant was only a cut above the enlisted man in technical knowledge, and he didn't know the answer either. But there is always (one hopes) a higher echelon of experience and ability when the phenomenon-finder — or his superior — can't quite see the light, and eventually the gold-bearing ore was assayed by one of the experts. A telegram was on its way within hours — with the beginning steps of solution immediately taken by the senders, of course.

There is a hint here also of the good cooperation which existed between field and home base. In this case, as in many others, exciting telegrams "crossed" in mid-air after a minimum period of hectic activity, so you could possibly have a mixed impression as to "who did what first." (Who really cares?)

One happy postlude to the above harmonious composition. The lowly enlisted man was made a Warrant Officer on the basis of his findings, and ended his career as a valued member of the establishment entrusted with post-war cryptography.

An equally impressive example of such persistence came about in my own shop, where a scholarly (weren't they all?) misfit "soldier" assigned to my technical staff doggedly pursued a trail in an IBM run.

He thought he saw a phenomenon, in one of many machine sorts we had made involving various combinations of possible "controls" for the indicator encipherment.

Without too much elaboration at this point, we knew such controls might be obtained by such devices as taking perhaps three digits, such as the first position of the first group, the first position of the second group, and the last position of the last group.

The possible combinations in such a search are enormous — even limiting ourselves to likely spots perhaps used before. As well as the first few and last few groups, we were even suspicious of such things as dates, group counts, and other sources for our magic numbers.

Now our tyro searcher had been looking at run after run, and happened to be looking at one which was sorted on the last digit of the group count, and two other positions — say the first digit of the third group and the last digit of the last group — giving a three-digit number. (The first five and last five groups of the message might be printed out, within the major sort of whatever control you were pursuing.)

When he came to the portion of the sort showing putative controls of 500, 501, 502 etc., he noticed certain blocks had an interesting consistency, For example, almost the entire set of message beginnings listed under control (?) 503 might show a consistent ABCB. ABBB, or AAAB pattern in the first groups of the messages — the second and the fourth digits being usually the same.

9383, or 4202 or 1999 or 7778 would be representative of the dozen or so beginnings listed However, there were a few mavericks present quite often — perhaps eight groups showing such repeats and a couple which didn't.

And only about ten percent of the hundred "controls" beginning with the digit 5 showed the property — even though it was undoubtedly significant in this repetitive property when it DID occur.

Doesn't sound like much, does it? And the relatively inexperienced GI hadn't the slightest explanation for what he thought he saw. But his calling it to our attention triggered the break-in of the indicator workings of a major system.

Why the digit "5" of message length, which marked the first digit of the control we happened to be looking at? Each text had a group count in the heading, such as "65" to show that there were exactly 45 4-digit groups in the message — normal procedure in many communications.

Since it was later proved that there were four non-textual groups, specifically placed to indicate the starting and ending points within a keying book, a message count ending in the digit "5" was effectively an indication of a real group count ending in "1" — 31, 41, 51, or 91 groups, for example.

In the special key book used to disguise starting and ending points of the specific additive used, there were 100 4-digit groups on each page — ten rows and ten columns, with scrambled numerical designations to identify them.

Thus, if the stretch of key did not run from page to page, but began and ended on the same page, any message of 21, 31, 41 groups, etc., ended with key from the same column as the first group.

Necessarily, any indicator (in its plain version) for such a message, which listed the numerical designation of the beginning row and column, as well as the designation for the ending row and column, would invariably have a repeated digit in the "proper" place.

It really wasn't as simple as it sounds, since the indicators were enciphered by a very complex system, which we will not go very deeply into at this time.

Suffice it to say that the fact that only 10 percent of the 500-control series had this property is due to the 10 percent chance that the additive key for that control also had a .B.B pattern.

Thus the seemingly unimportant discovery of "interesting" repeats in limited and specific cases provided vital road signs. It showed where the indicator was, what the various positions were intended to indicate, and the "control" for its encipherment.

And it is perhaps obvious that when an indicator system for many problems is solved — including machine ciphers — an important battle in the cryptanalytic war is won.

Another happy ending — the scholarly GI, who started in my section as a corporal, breezed through OCS to become one of the outstanding lieutenants on my staff — although it didn't require the minor victory I have mentioned to prod us into pushing for his gold bars.

Chapter XIII

MISCELLANEOUS PROBLEMS

FINALLY, I'll throw in a few problems which could be of some interest, even though some of them are specialized cases which only certain solvers might recognize. Also, we might search for a few examples of every-day problems, or posers which one meets in recreational pursuits, or "newspaper" puzzles which may be easier to solve if one approaches them with a diagnostic eye.

Railfence Transposition

The plaintext is written in zigzag, and the ciphertext taken off by rows.

```
C   L   F   D   H
 I I W R I L C P E
  V   A   E   I   R
```

Ciphertext: CLFDH IIWRI LCPEV AEIR

One usually confronts keys of five or six "rails" (rarely seven) rather than the three in the example. It is important to remember that the top and bottom rails are "half-size" and (approximately) as a pair equal the length of an ordinary rail. Let us (together) solve a problem which has a "tip" of GLAD.

LRLDB BIIGA GOAVE EOEFT FANDR SERGT POESD NIUNU LDATI AOTAR

Although there is a tougher version of this system which takes the rows off in some keyed order, the simpler version here may have the letters G L A D in some sensible positions in the cipher — either some G coming before some L before some A before some D, or (depending on which row the G appears) a permutation such as L A D G — going around the corner, so to speak.

One can note the distances between occurrences — such as

```
9 11 29   1 3 41   10 13 22 43 46 49   4 24 35 42
G         L        A                    D
```

and (bearing in mind that one may have to circle "around the corner") look for intervals that would make sense in a 50-wide cipher. I see a couple of interesting differences (and would be happy to get a couple of consecutive 10's for example.)

But there may be an even nicer thing to look for — in the case of the crib running over the top or bottom of the "fence," so that GA or LD may appear together, with the in-between letter L or A, as the case may be, on the top or bottom row.

Since the top and bottom lines have roughly half as many letters as the main rows of the "fence" one cannot simply assume a width of five rails and write the above text out in sections of ten letters each, then try to line things up. We might have a case where the first "clump" will be ten letters, followed by a clump of five, then another five, and finally two more tens.

It may not be all that confusing with a favorable text count (such as the 50 given) but if we are given some length where the text runs out after 68 letters or something, it poses more of a jigsaw.

In any event, the problem (as you would readily conclude) consists of lining up the rails on the basis of the crib, or good-looking plaintext. In the case given, there are two spots where LD occurs, and one showing GA, which should be tried (remembering that neither one may be valid, and that GLAD goes down — or up — without any doubling-back.

```
        D                                   G
B I I G A G O A V E   or   U N U L D A T I A
      L                                   A

                    or

        G
L R L D B B I I
      A
```

There are 5 Ds, 3 Gs, 3 Ls and 6 As that could to be tried in the above spots, hoping that some glimmer of plaintext would show up. The correct placement will "make sense" both in plain context and arithmetically, since it should divide the cipher into segments that fit a presumed width.

You can play around with the possibilities, but the one I hit on is:

```
F A N D R S E R G T P O
E S D N I U N U L D A T
  I   A   O   T   A   R
```

You will note that the F is the 21st letter, the E is the 33rd, and the I is the 45th, giving 12 letters in the two complete rails, and finishing nicely with six for the "short" rail.

I'm sure you can easily spot the words which will automatically give you the complete text.

Swagman Transposition

A fairly easy form of transposition, which doesn't involve quite so much trial and error as some systems, is the Swagman. (I don't know, but suspect that the name is an offshoot of the ordinary "Tramp" which some cryppies use to delineate Transposition.)

One needs a keying square composed of some random (?) arrangement of digits (between 4 and 8 wide) wherein the digits have no repeats within either row or column. For example:

```
3 2 1 4 5
1 5 3 2 4
2 4 5 3 1
5 3 4 1 2
4 1 2 5 3
```

The plain text is written horizontally to form a rectangle commensurate with the width selected — here 5 x 5. If necessary. nulls are used to complete the the last letters of the rectangle, which must be a multiple of the width of the square.

Suppose we have the short "message" and the corresponding 4-wide box:

```
T H I S   I S A N   E X       2 3 1 4
A M P L   E O F T   H E       3 2 4 1
S W A G   M A N T   R A       4 1 3 2
N S P O   S I T I   O N       1 4 2 3
```

The first column of each box is rearranged according to the first column of the square — thus T is put on the second row (as I and E of the other boxes will be) A is put on the third, S on the fourth and N on the first. Then H (and S and X) get put on the third row, M on the second, W on the first, and S on the fourth — since the second column of our digital box is 3 2 1 4. And so on, to form an intermediate version:

```
N W I L   S A A T   O A
T M P G   I O T T   E E
A H A O   E S N I   H X
S S P S   M I F N   R N
```

The final cipher version is taken off vertically:

```
N T A S W   M H S I P   A P L G O   S S I E M   A O S I A
T N F T T   I N O E H   R A E X N
```

Let's look at a problem:

```
POGTC VEEIO AIROR LLDLE NOWGP AIAAN FNGTA THATL ICTPN HUEAX
YGELA DIDAN EUNMB ILANT RRICM EAMIG LAMPA RTASR POOOA LUPDO
BROAS ESELA NSNQL ODUHC EIAAS CGDSO ORREM BTOWI SOUS
```

The crib given is Caesar UXPFMP, which I'm sure you now would decide must be TWOELO, since all the rest of the possibilities are even stranger. (It might tip us off that the text could be equally strange.)

```
POGTC VEEIO AIROR LLDLE NOWGP AIAAN FNGTA THATL ICTPN HUEAX
 o     ee          ll l   o
YGELA DIDAN EUNMB ILANT RRICM EAMIG LAMPA RTASR POOOA LUPDO
```

```
BROAS ESELA NSNQL ODUHC EIAAS CGDSO ORREM BTOWI SOUS
                                          T W   o
```

Since there are only a couple of W's, we look for the other letters of the crib IN ORDER around that area, and find them, albeit ambiguously. Now, we must come up with a width which puts each letter in consecutive columns. The unique TW and the final O should narrow things down.

The break must be at position 138, or one position removed. 138 is a factor of 6, which looks reasonable, since the total count of 144 factors to that number. (Since 4 and 8 and 9 are also factors, we did the right thing in looking for the crib first.)

And when we mark that area out, we find we can put one letter of the crib in each of several consecutive boxes, so we might as well prepare the worksheet by writing the cipher down the 6-deep columns.

```
P E R L P F    H T A A E L    I I A P U O    L L E G R W
O E O E A N    A P X D U A    C G R O P A    A O I D E I
G I R N I G    T N Y I N N    M L T O D S    N D A S M S
T O L O A T    L H G D M T    E A A O I E    S U A O B O
C A L W A A    I U E A B R    A M S A B S    N H S O T U
V I D G N T    C E L N I R    M P R L R E    Q C C R O S

   Plain:

        A F           B L           P O           T W
O E L O n g    A T E D i n    C I S O r s    A L S O o s

   Key:

. . . . 5 1
2 1 5 4 . .
```

You will see that when we break the text into sets of six groups of six letters, one row of the digital key must be 2154?? (the ambiguities disappear) and key row just above must be ????51. We haven't yet determined which rows these are, but as we go through the message six times picking up the various levels, everything should come out in the wash. Putting in the 3/6 possibilities on the 2 1 5 4 row, one sees the obvious text "two elongated incisors also os." I am sure you can prepare your own worksheet and read the entire text, and recover the 6 x 6 square.[1]

Homophonic

This system, though not particularly difficult to solve, even without a crib, is in some ways more like traditional substitution systems.

It merely uses four 25-letter normal alphabets, each assigned to the four divisions of the 100 dinomes, 1-25, 26 to 50, etc. However, each alphabet starts in some prearranged spot within the 25 dinome range, running cyclically "around the corner" so that if A is 24, B is 25 and C is 01, D 02, etc. (00 is the end of the range beginning with 75.) In accordance with the usual procedure, I and J occupy the same position.

A four-letter keyword may be called for, reading the positions 01, 26, 51 and 76.

Given a crib, the job is fairly elementary, and you would probably figure out how to tackle a problem even without a crib. We'll give you one to cut your Homophonic teeth on. It's so easy that I'll not finish the long problem they gave in the Analysts Corner and I'll give you just enough to get started on a solution.

```
23 26 59 93 12 36 64 92 19 46 57 93 46 64 95 12 46 63 26
95 95 12 92 94 05 32 94 64 93 12 36 64 92 19 46 57 93 46
64 95 32 57 57 26 95 53 63 12 92 94 85 94 93 59 94 64 93
12 36 64 92 19 46 57 93 46 64 95 32 57 57 26 95 86 26 94
05 19 32 63 64 93 12 36 64 92 19 46 57 93 19 12 64 95 12
63 05 26 53 26 95 63 46 93 12 36 64 92 19 46 57 93 46 64
etc.
```

Frankly, it is easy to get started without a crib; what you might look for is an area where several dinomes in the same "quadrant" are together, or at least nearby.

Look at the area 95 95 12 92 94. All you have to do is "complete the plain component" of the four high numbers (remembering to skip J, and going around the corner.) You can start with any arbitrary value — say A for 95. To save space I will extend horizontally, but you may want to go down.

(You could also take a frequency count of the entire message, since you would see a typical normal-minus-J frequency looping around each quadrant.)

```
95   A B C D E F G H I K L M N O P Q R S T U V W X Y Z
95   A B C D E F G H I K L M N O P Q R S T U V W X Y Z
12
92   X Y Z A B C D E F G H I K L M N O P Q R S T U V W
94   Z A B C D E F G H I K L M N O P Q R S T U V W X Y
```

MM.IL and SS.PR are possible, but the FF.CE looks promising for guessing (OFFICE? FACE?) — so let's assume that is right, and fill in all the values for the quadrant 76-00. While we're at it, let's assume the O (in small letters):

```
26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50
o p q r s t u v w x y z a b c d e f g h i k l m n

76 77 78 79 80 81 82 83 84 85 86 87 88 89 90
 M  N  O  P  Q  R  S  T  U  V  W  X  Y  Z  A

91 92 93 94 95 96 97 98 99 00
 B  C  D  E  F  G  H  I  K  L
```

```
23 26 59 93 12 36 64 92 19 46 57 93 46 64 95 12 46 63 26
   o     D     y     C     i     D  i     F     i     o
95 95 12 92 94 05 32 94 64 93 12 36 64 92 19 46 57 93 46
 F  F     C  E     u  E     D     y     C     i     D  i
64 95 32 57 57 26 95 53 63 12 92 94 86 94 93 59 94 64 93
   F  u        o  F           C  E  W  E  D     E     D
12 36 64 92 19 46 57 93 46 64 95 32 57 57 26 95 86 26 94
   y     C     i     D  i     F  u        o  F  W  o  E
05 19 32 63 64 93 12 36 64 92 19 46 57 93 19 12 64 95 12
      u        D     y     C     i     D           F
63 05 26 53 26 95 63 46 93 12 36 64 92 19 46 57 93 46 64
      o     o  F     i  D     y     C     i     D  i
```

The rhythm and many repeats in the problem suggest that the composer was quoting some poetry, perhaps — but I'm sure you see the oldie represented, and could fill out the whole thing.

Grandpre

This is kinda fun, and can be fairly difficult. An 8 x 8 square is prepared, with eight selected words written in the rows in such a manner that the first column is an acrostic — spelling another eight-letter word. Common sense dictates that the composer choose words which have among them all the letters of the alphabet, so you can expect QUAGMIRE, or TAXICABS or JONQUILS (the last one containing a couple of low-frequency but necessary letters.)

The square I am looking at has LACQUERS down column 1, with the eight rows being LADYBUGS, AZIMUTHS, CALFSKIN, QUACKISH, UNJOVIAL, EVULSION, ROWDYISM and SEXTUPLY — hardly a lexicon of household words, but sufficient unto the day.

Naturally, the square has coordinates 1-8 at top and side.

```
   1 2 3 4 5 6 7 8
1  L A D Y B U G S
2  A Z I M U T H S
3  C A L F S K I N   etc.
```

The plain they show as an example, and the numbers selected (there being a choice for common letters, of course) is:

```
T  H  E  F  I  R  S  T  C  O  L  U  M  N
84 27 82 34 56 71 77 26 44 54 64 63 78 52
```

Let's look at a problem. The tip is THROUGHTHE.

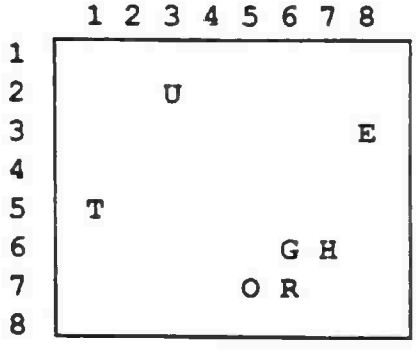
```
35 51 67 12 44 83 38 38 22 88 53 61 48 26 17 15 34 53 78 28
22 64 45 63 64 78 26 84 47 82 11 75 38 43 51 67 76 75 24 66
67 51 67 38 56 71 27 85 75 45 65 81 11 88 34 53 11 82 88 64
81 26 52 84 25 61 12 33 74 16 46 31 14 85 65 16 53 51 65 86
78 48 61 68 35 85 58 37 25 61 75 78 31 64 81 26 25 38 15 76
16 52 56 71 76 51 67 64 11 57 61 85 26 32 41 21 62 65 44 52
74 82 81 51 77 56 26 52 64 35 54 78 86 16 38 62 51 53 54 48
48 38 16 71 47 42 51 65 71 22 77 56 51 67 46 68 38 45 52 71
48 61 48 81 77 51 17 61 78 11
```

Our crib has repeats TH....HTH — can we find any spot in the message which would accommodate such, if the composer happened to choose the same number even though variants are possible? Notice the stretch 51 67 76 75 24 66 67 51 67 38. A small point, but you also did not get conflicts caused by a repeated digit which falls under a different plain — a patent impossibility.

Let's start a square and fill in all the values we can:

```
35 51 67 12 44 83 38 38 22 88 53 61 48 26 17 15 34 53 78 28
    T H             E  E
22 64 45 63 64 78 26 84 47 82 11 75 38 43 51 67 76 75 24 66
                                  O        T  H  R  O  U  G
67 51 67 38 56 71 27 85 75 45 65 81 11 88 34 53 11 82 88 64
 H  T  H  E              O
81 26 52 84 25 61 12 33 74 16 46 31 14 85 65 16 53 51 65 86
```

```
78 48 61 68 35 85 58 37 25 61 75 78 31 64 81 26 25 38 15 76
                                  O                  E      R
16 52 56 71 76 51 67 64 11 57 61 85 26 32 41 21 62 65 44 52
               R     H
74 82 81 51 77 56 26 52 64 35 54 78 86 16 38 62 51 53 54 48
            T                                       T
48 38 16 71 47 42 51 65 71 22 77 56 51 67 46 68 38 45 52 71
    E                T                  T
48 61 48 81 77 51 17 61 78 11
                  T
```



Now comes the tough part — making some shrewd guesses as to either the text or the words on the square. Since the latter doesn't seem to offer a wedge right now, an inspired guess might be that the message starts "It has been said that" — since the double E hints somewhat in that direction. Fill 'er in — small letters, of course. (You will note the encouraging positions of S and D in the 8th column of the square — and although we see no obvious words, there are no improbable combinations.)

```
35 51 67 12 44 83 38 38 22 88 53 61 48 26 17 15 34 53 78 28
 i  T H   a  s  b  E  E  n  s  a  i  d  t  h  a  t  a
22 64 45 63 64 78 26 84 47 82 11 75 38 43 51 67 76 75 24 66
 n                 t                 O        T  H  R  O  U  G
67 51 67 38 56 71 27 85 75 45 65 81 11 88 34 53 11 82 88 64
 H  T  H  E                 O                 s  t           s
81 26 52 84 25 61 12 33 74 16 46 31 14 85 65 16 53 51 65 86
    t                 a                             a
78 48 61 68 35 85 58 37 25 61 75 78 31 64 81 26 25 38 15 76
    d  i     i                 i  O                 t     E  a  R
16 52 56 71 76 51 67 64 11 57 61 85 26 32 41 21 62 65 44 52
               R     H                 t                 s
74 82 81 51 77 56 26 52 64 35 54 78 86 16 38 62 51 53 54 48
            T        t        i                 E     T  a
48 38 16 71 47 42 51 65 71 22 77 56 51 67 46 68 38 45 52 71
 d  E                T        n        T
48 61 48 81 77 51 17 61 78 11
 d     d        T  h  i
```

	1	2	3	4	5	6	7	8
1		a			a		h	
2		n		U		t		
3				t	i			E
4				s				d
5	T		a					
6	i					G	H	
7					O	R		
8			b					s

78 certainly looks like an N, with the IO before it. The end of the message now looks like noTHING.

We have to jump in someplace, and (to me) the row 6 of the square looks like a good place, since the GH offers only a few words, such as INVEIGHS. (Computer buffs or those with certain pattern lists merely look for I——GH-)

The area in the text which runs 64 45 63 64 right after the text we have already established clinches the deal —but you should now finish this up yourself.

Pollux

Two systems which (like Fractionated Morse) utilize dots and dashes in the enciphering process are Pollux and Morbit. Since Pollux is merely a substitution on the three individual symbols for dot, dash and space (x), it is rather simple and straightforward. It uses the 10 digits, giving room for three variants for two symbols, and four variants for another — usually the dot.

(You will remember that Fractionated Morse used 26 of the 27 three-symbol elements possible, from ••• to **xx•**, assigning each one a letter, using a keyword mixed alphabet. Morbit is in the middle of complexity, since it uses nine selected letters to accommodate the possible combination of elements two at a time — from •• to **xx-**.)

With regard to Pollux, a crib would make things almost too easy, but without a tip it still does not pose an insoluble problem. I personally find that making up your own crib, such as "THE" will work almost as well, since the Morse string of symbols is long enough to throw out most positions if you slide the "crib" along, looking for confirmations and throwing out slides in which a digit 4, for example, has to be an X in one place and a dash in another.

But let's see what we can do without a crib or a "guessed word." Take the following message:

```
08639 34257 02417 68593 30414 56234 90774 56669
12446 50379 71547 91676 05349 89256 20394 15303
98481 60537 46938 34154 09343 55209 31060 29394
86150 79781 34540 49460
```

My analysis may sound somewhat like a Logic problem, in that we can write down several statements that are true, and list several "either or situations."

1. 0 must be dot or dash, since (at least to most constructors) one doesn't start out with an X.

2. We are fortunate in finding three digits repeated as a trigraph — 666 at the end of the first line. That digit cannot be X, since we cannot have three X's in a row. (By prearrangement, one could use three X's as a period, or sentence separator.) By the same token, 5 or 9 MUST be an X, and perhaps both are.

3. Knowing that both 0 and 6 must be dot or dash, we can see if that yields new information. Note the string 65520. It is obvious that either 5 or 2 must be X, or perhaps both. But we find an occurrence of 552, which says that one of them must be a non-X.

4. Looking for ABA patterns, we note several: 393, 797, 989, 848. 383, 454, 343, 494, 797 etc. It is worth noting such, because one of these digits in each case must be a non-X.

5. At the beginning, 8 6 3 or 9 must be an X, since no Morse letter is more than four symbols long. (If numbers are used, which are 5 symbols, they usually tell you.)

I could go on in this "logical" fashion, and start making either/or charts, etc., and eventually narrow things down. But my instincts are beginning to tell me that the way the digit 9 behaves in the cases I have so far listed makes it worth an assumption that 9 is an X.

How does it look if I try 9 for X? For one thing, 3, 7 and 4 must be non-X, as well as 0 and 6. I immediately discover that 5 must also be X, because of the occurrence of 934355. I will list the known "dot or dash" numbers as # below.

```
0863934257024176859330414562349077456669
####x###x####x###xx####x#x####x####x###x

1244650379715479167605349892562039415303
x#####x###x#xx##xx###3x##x#x#x####x#xx###

9848160537469383415409343552093106029394
x###x##x####x####xx##x###xx##x#x####x#x#

86150797813454049460
##xx##x##x##x###x###
```

Adding in the 5X assumption above, we see that we had to add 8 and 2 to the non-X category. And the 570241 stretch (x# ## #) then tells us that 0 or 1 must be X — but we decided the message should not start with X. So — we can write out the numbers and substitute X for 1,5 and 9, and think about a good plain text guess.

```
0863 342 7024 768/ 3304 4 6234 0774 666/
2446 037 7/ 47/6760 34 8 2 6203 4 / 303
848 60 3746 3834/ 40 343/20 3 0602 3 486/
07 78 34 404 460
```

When we examine the possibilities, we are inclined to think that 6 should be dot, since S is a much better ending of a word than O. And the single word 47 might more logically be A than I — we can back track if necessary.

```
                  0 1 2 3 4 5 6 7 8 9
                    x     . x .     x

0863 342 7024 768/ 3304 4 6234 0774 666/
  .   .   -   . -.       . . .  .  --. ...
                            E       p?  S

2446 037 7/ 47/6760 34 8 2 6203 4 303
 ...            .-.   .     .    .

848 60 3746 3834/ 40 343/20 3 0602 3 486/
 .   .  -..      .   .   .        .     . .

07 78 34 404 460
 - -    . . . ..
m? n?  i? r? s?
a? m?     s? r?
```

I am sure you would put in BUT and AMISS, and clean up the entire message.

Now I want to show you an alternative message very similar to the one you just solved, but because of the way I chose to encipher it, the approach we used might fail. However (remember the swings and roundabouts) there is an intriguing phenomenon which I wonder whether or not you can spot.

```
12346 50379 81542 91670 35469 81250 73496 15034
98621 03547 60938 46150 39460 15239 41603 75496
10153 29781 46503 49603 15293 46013 59460 71827
53498 62013 59782 14670 15398 46213 78456 90231
46509 73842 65091
```

Try your hand at discovering "something funny" and only resort to the explanation in the Notes if you really "can't see a thing."[2]

Morbit

Let's look at the closely related Morbit system.

Choosing a 9-letter keyword, which yields an order for the 9 digits by calling the lowest letter 1, and so on, the two-element combinations are laid out:

```
W I S E C R A C K
9 5 8 4 2 7 1 3 6
. . . - - - X X X
. - X . - X . - X
```

Enciphering "Once upon a time" we would get as Morse equivalents - - - X - • X - • - • X • X X • • - X • - - • X - - - X - • X X • - X X - X • • X - - X • X X which would be broken up in pairs, and enciphered as 2 7 4 3 5 8 8 1 5 1 2, etc.

Here's a message — with no crib given.

```
47673 46678 32626 71644 31327 16449 28351 62636 29544
39413 51432 66289 24376 74685 14794 76357 16851 62332
98524 34716 45166 35274 91326 85239 54437 15346 13456
31647 73749 19461 43524 98542 98616 17647 39468 57232
35289 17626 34383 37913 26471 66326 43349 45274 33276
61734 67346 63477 8766
```

A few obvious statements. Any doubled digit (44, for example) cannot be XX or - - although it could be • • (H being • • • •) The first value, 4 in this case, cannot be X-anything.

There are a few such helps, but the simplest entry is perhaps made by sliding a logical word, such as THE through the text. Write down the Morse equivalent to your "crib" on a strip of paper, using the same size squares as your worksheet. However, you will have two possible versions, starting with either xx -x •• •• x• xx or ?x x- x• - - •• •x •x x?

You will have something that looks:

```
4 7 6 7 3 4 6 6 7 8 3 2 6 2 6 7 1 6 4 4 3 1 3 2 7
-x....x   immediate conflict (and you can skip a few)

xx-x....x.xx conflict on 6
 xx-x....x.xx confirmation on 6 — check further
  xx-x....x.xx confirmation on 7, conflict on 6
   xx-x....x.xx  conflict on 6
    xx-x....x.xx  confirmation on 6, check further
```

You will note that the last case above implies 3 is xx. Are there any 33's? Sure enough, at the end of the second line you will see a 33, and one can't have xxxx.

Let's really try the placement that had no conflicts, above.

```
0 1 2 3 4 5 6 7 8 9
      x .   . x
      . .   x -
```

```
4 7 6 7 3 4 6 6 7 8 3 2 6 2 6 7 1 6 4 4 3 1 3 2 7 1 6
. . x - . xx - x . . . . x . xx -   x .   . x   . xx -   . x . . . x   .   x .   x -   . x
  IN/THE/ H
```

Could we find anything better than a start of IN THE? The H a little further on should be preceded by a 4-element letter, or else go 124, or 214 elements. That narrows things down, since the best 4 element letter before the H would be C and we have the dot and dash correctly placed already. So —

```
0 1 2 3 4 5 6 7 8 9
  .   x .   . x
  -   . .   x -
```

```
4 7 6 7 3 4 6 6 7 8 3 2 6 2 6 7 1 6 4 4 3 1 3 2 7 1 6 4 4 9
. . x - . xx - x . . . . x . xx -   x .   . x   . xx - . - . x . . . x . . - x .   x - . - . x . . . .
    IN/    THE /                          /CHURCH?
```

I leave the rest to you.

Amsco

Some systems are fairly easy to describe, and the techniques of solution fairly obvious. I will include a few, with a minimum of "how to" — since by now you are undoubtedly willing and able to say to yourself "let's see. What do I have to do to reverse the enciphering process?"

Amsco is an incomplete columnar transposition, with an added fillip. The plain is inscribed in usual fashion, but in taking off the columns (according to some key) one alternates single letters and digraphs. You might want to use cross-section paper with large enough squares to make entry of digraphs easy.

4	1	3	2	5
IN	C	OM	P	LE
T	EC	O	LU	M
NA	R	WI	T	HA
L	TE	R	NA	T
IN	G	SI	N	GL
E	LE	T	TE	R
SA	N	DD	I	GR
A	PH	S		

Cipher: CECRT EGLEN PHPLU TNANT EIOMO WIRSI TDDSI NTNAL INESA ALEMH ATGLR GR

If you have a crib, obviously you must find about half of the letters in it as digraphs in the cipher. It is especially helpful if the crib is a bit wider than the transposition key, so that you may be lucky enough to establish width quickly.

In the above example, suppose you were given the ..TINGSINGLE.. The G would be a good letter to look for, since it is not too common a plain, and we have two G's in the crib.

Do we find an interesting spot? We do — in fact there are two places to investigate, since we have two occurrences of GL and one of them has a T in front (which might, though not necessarily, indicate a width of 5 — the plain being T IN G SI N GL E) and the other GL could explain no "fore or aft" hit by a longer key.

There are no NG's in the cipher, but remember that the G could be a single letter, in the column C EC R TE G LE (which works fine at an agreeable width of five) or one of the G's in L EM H AT G LR G R(ending)

You could work out your own procedure of looking for good columnar associations, just as you know how to do in ordinary columnar transposition.

See if the following falls to pieces just as easily. (The tip is WONDERFUL.)

```
RVEEW ISELO WOETH HIWTA HEETH HYORH OTEAL ITTAU SERRK
RETOL SORSI CUNNO ULAVE REOVO RBEEG EFFUG IMNHM CILEO
FSINW EEELI FOYFP LBFIS TDEWI RP
```

Myszkowski

This system (don't ask me where such names come from.) is also a twist on the usual columnar transposition, in that a pattern keyword (such as CANADIAN — which has several repeated letters) is used. Unlike the usual procedure, where such a word would result in a key 4 1 7 2 5 6 3 8, the repeated letters get the same number — 2 1 5 1 3 4 1 5 in this case.

Text is written in the box horizontally, in the usual fashion, but taken out by reading on each row all the same numbers, 2 or more at a time if necessary — in this case 3 letters under the 1 on the first row, then three letters on the second row, etc. And example will make this clear.

```
B A N A N A
2 1 3 1 3 1   Cipher text:  NOPEE OUNRI
I N C O M P   HATRW RKYNL TESNE SMNME
L E T E C O   TKNFA RSILL WTOAT DEROO
L U M N A R   CMTCM ATPEN DEDER URAUB
W I T H P A   AEFCS
T T E R N W
O R D K E Y
A N D L E T
T E R S U N
D E R S A M
E N U M B E
R T A K E N
O F F A C R
O S S
```

Given a crib, the approach is similar to that used in Amsco, above, but instead of adjacent letters of a plain digraph, one must search for a string of, for example, three letters in the right order and adjacent on a row, but not necessarily adjacent in the plain.

Thus, if given COLUMNAR as a tip, one would hope to find at least a pair of letters, or possibly three, in the right order such as OMR, or LUA. In the example above, you will see U N R — and the O in front of it might be the proper O to give you a width of six. (It might not be the right O, remember, if the width is longer than the crib.)

Writing out the text on this basis, one has an approximation of the box —

```
N   O   P
E   E C O L
U M N A R
I   H   A
T   R   W  etc.
```

You don't immediately know the exact arrangement, since N could be the first or second plain, but you know the width is six, and hence the depth is three columns of 13 and three of 12 letters.

Looking for CA, you might find it as a digraph CA, or as CLA, or CMA even the unlikely CLMA. Reflect a little, and you will see why this is so in the very fortunate example you were given. You find it as CMA, which gives you most of INCOMPLETE for a beginning, when you put in the CMT ahead of the CMA.

The stream does not always flow so smoothly, but you should be able to assemble columns on logical plaintext, just as you do with ordinary columnar transposition — and the AMSCO you just met. If you have any idea of the text, you might dream up a word longer than the probable width, such as TRANSPOSITION, and look for such "skip patterns" as TAII.

Better still, check all the cipher P's and hope for one or more of the letters TRANS ahead of it and the confirming letters taken from OSITION, such as TAPSIN together — which admittedly is almost too much to ask for.

Nihilist Substitution

Historically, the Nihilists evidently used at least two cipher systems, neither one of which should have given the State Police (or whoever) much trouble, if they had any cryptanalytic experience. The Nihilist Transposition ends up as being merely a columnar transpo in which lines do not necessarily follow one after another, but are taken off by a key. (A nuisance to follow but not much of an obstacle.)

The Substitution is a bit more fun and a bit tougher, but by now I am sure you would spot the phenomena and figure out a method of attack.

A 5 x 5 square is used to set up the plain text (in dinomes) on some predetermined cycle — say seven dinomes wide. Then each column of the seven dinomes is enciphered by the appropriate one of seven key letters, which are also put through the 5 x 5 square.

Effectively, this creates seven two-digit key values which are added to the plain dinomes in each column. The enciphered values are then read off horizontally to create the message. (The Nihilists used "carrying" in the few cases where 5 and 5 are added, but realized they had to change the numbers from 100 to 110, making them 00 to 10.)

An abbreviated example:

```
          Key:   Z  E  R  O
(using square)   55 31 43 15
                 ---------

Plain            T  H  I  S
                 42 33 12 23
  CIPHER:        97 64 55 38

                 I  S  A  N
                 12 23 13 25
                 67 54 56 40

Cipher 97 64 55 38 67 54 56 40 etc.
```

Obviously the first single column of digits in the example above could be only 6,7,8,9 or 0, and the second column as well. But the third column (under the first digit of E) could be only 4 to 8 and the next column 2 to 6. This makes the determination of the width a minor chore rather than a brilliant deduction — unless one has never heard of the Nihilist Substitution. (Note that "carrying" addition is used, converting 25 plus 15 to 40.)

In such case, you would probably line the message up on the right cycle on the basis of some nice repeat, and then discover the limitation of five digits within any one column. Relatively easy?

Naturally a crib would help you get started in recovering the square, but a few logical guesses based on frequency might suffice, since we really have a simple substitution to contend with.

One for you to solve:

```
67536 67956 69376 48456 56474 73385 69747 94963 95764
55846 35957 67766 58456 65774 69493 58285 48485 82776
59544 04945 76898 44826 45758 57658 55350 36774 50393
59276 48464 54566 79664 93537 66766 66939 45945 74746
28639 46748 58465 57488 65465 84574 56486 0
```

After you find the cycle (perhaps checking the interval between occurrences of 0 — since those columns have to be 6 7 8 9 0 and are easy to pin down) try to "strip the key." You may have a few ambiguities, with only 4 digits in a column, or something.

But starting out the text with WITH should get you on your way.

Final Comments

The only thing left in our Problem Solving effort is a series of challenges for you to tackle. Chapter Fourteen includes several brain-teasers, some of which are based on cryptographic systems and some of which are simply chances for you to do a little diagnosis or statistical evaluation.

Chapter Fifteen will provide a why and wherefore, in case the phenomenon you may have found in Chapter Fourteen leads you nowhere — or perhaps it will be a case of "Oh, I should have seen that."

But before we take pencil in hand for the "final examination" it might be fitting to close the expository section with a few remarks and reminders on *Problem Solving*.

Several "buzzwords" (if we may borrow such frowned-upon appellations from the modern business world) might be worth keeping at eye-level in your inner sanctum, or computer room, or wherever you sit down with cross-section paper or dictionaries or puzzle books. One phrase (admittedly it's not a buzzWORD) I would like to see engraved is DON'T IMMEDIATELY SAY, "I DON'T KNOW."

While the idea of honest admission of ignorance may be commendable in certain cases, I find that some people use the "I don't know" expression perhaps even before they have grasped the significance of the question. As an automatic response, it should be replaced with at least a "Let's see."

As a second wall-hanger, I would replace the "THINK" card (or THIMK, if you prefer) with A REALLY worn cliche — "PRACTICE MAKES PERFECT." For almost

any skill will deteriorate if you don't use it, and the converse is usually true, that (to a reasonable point) the more you exercise the stronger you get.

This means that if you really expect to master the justly Cryptic Crossword, for example, you can't just wait until you make that trip to London, where a TIMES is available. Many puzzlers tell me that the fifth or sixth one of my *The Nation* offerings they confront is much easier than the second or third, and that they get in the swing of things readily as they begin to personalize the battle — "Getting to Know You" being the theme song, perhaps.

But many papers and magazines are beginning to run Cryptics — some of them rather good, and some pretty awful — so you can make some attempt to hone your skills rather than let them get rusty.

If you skip some supposedly solvable problems merely because you consider them "hard" they will only become harder. The statistics show that many more solvers do the Aristocrats than do the Patristocrats — which implies some fear of taking that next step.

And so it is with many of the skills you may use in day-to-day problem solving. Is "memory" a skill? Perhaps you are blessed with a good one, or cursed with a bad one, due to genes — but the utilization of what memory you have is certainly a skill that can be improved, or can deteriorate.

So, remembering that PRACTICE MAKES PERFECT, I suggest that you use whatever techniques you can to conscientiously improve your memory, instead of bemoaning its demise.

And please don't kid around with the "I must be getting old" bit. While it seems to be true that certain age-related diseases play havoc with memory (and I obvously am no medical expert) there is no reason to believe that for the majority of people a "good" memory cannot be sustained — and even improved — far beyond retirement age.

I am no advocate of the numerous "memory courses" one sees advertised on TV, or tapes you play near your pillow while you are asleep, or artificial association techniques that lead you by way of a three-digit number to the fact that Mr. Jones is a plumber. (Remember that certain "association techniques" are perfectly valid.) But you don't have to part with a single dollar to set up your own regimen of memory development.

Perhaps my second theme song would be "Try To Remember" — since one may have to make a conscious effort, whether it be "remembering" the name of the movie star, or finding the answer to "where did I put my glasses?" Part of such "trying" might be even the rudimentary process of going quickly through the alphabet (mentally, of course, unless you don't mind your ABC's being made public) especially when trying to think of a name, or a word.

When you come to the right letter, you may be amazed to discover some inner feeling that you are getting close, and suddenly the name is there. I personally (again as a layman) think it astounding that I will say "it's a two-syllable word — six or seven

letters — perhaps begins with M" when I really haven't the foggiest idea of why I said it. (The answer will probably be Marquand, because QU sounds like one letter.)

Like the computer I am now working with, your brain seems to have that capacity of "random searching" and you may be surprised at the number of times it finds something you were looking for, even without very specific directions. I often find that if I am trying to recall something (the name of that movie star, again?) instead of beating the problem to death, or intensive searching, if I sort of nudge my memory "computer" and then turn away from the dilemma, all of a sudden the memory screen sees the name pop up.

Another type of "practice" (or exercise) hinges on the fact (as my lay mind remembers it) that short-term memory, especially, must be reinforced if you expect impressions to endure. Therefore, as a result of a suggestion made to my young daughter twenty years ago, while on a guided tour of Morocco, I can still "see" many of the country's several capitals.

I passed on to Wendy my trick of trying to recapture some of each day's adventures by "thinking" about them before drifting off to sleep. Today, I can still give you a rough description of where we broke our journey, sometimes "seeing" the hotel lobby (perhaps not the name) or "seeing" the broad avenue (from the correct floor window, and the correct side of the street) or "almost seeing" the eccentric uranium miner from "Down Under" who sat in front of our back-window seats on the bus.

I agree — most of this is trivia, and not really memorable. But I don't think it overloads your (very capacious) memory to have a few things retained that are "pleasant" rather than "important."

My point is that the conscious effort, or practice, or exercise, or whatever you choose to call it, may be one of the reasons for my feeling that my memory (at my relatively advanced age) is perhaps better, rather than worse, compared to my more greenish years when I didn't expend much energy in trying to keep the machinery up to scratch.

It goes without saying that if I abandoned my hobbies of creating and solving crosswords, playing bridge, having fun with cryptanalysis (perhaps even writing a bit about the good old days) and maintaining at least an acquaintance with the many dictionaries, reference material, and "particular interest" books in my computer room, I could easily slip into the attitude I decry in some of my friends and neighbors — "it's because I'm getting old."

There are many opportunities in your every-day life to "practice" some of the desired skills. Perhaps you play cards. If poker, you have a beautiful opportunity for playing around with simple probabilities — as long as it doesn't slow down your game or interfere with your usual native savvy.

If bridge is your meat, it doesn't mess up things at all to calculate your chance of dealing after three others have exposed their cards — assuming you don't show off with a professorial exposition of probabilities.

You possibly marvel at good players who can seem almost able to tell what cards an individual possesses, or can dredge up a complete hand when the tournament director

asks for help in straightening out a messed-up board at duplicate. It takes effort and practice (assuming a certain amount of ability) but you, too, could be a lot better at such things than you are, or think you could be.

If your bridge bidding system utilizes point count, as I'm sure it should, do you automatically try to size up the approximate count in your partner's holdings, as well as the total between you? (Is it enough for game, or slam, or should you back off?)

Many people play Scrabble (and similar word games) almost blindly, seeming to think that analysis of any situation is making too much of a good thing out of what should be a pastime. They expend blanks or S's or wild cards (whatever the game calls for) with little regard to "am I wasting my assets" or will put themselves in a situation where only a drawn K would help them.

Of course not everyone "wastes time" in such trivial pursuits — and T. P. may be just the place to practice dredging up names, places and such. But even problems at the wood-working bench, or the computer, or even the kitchen stove may have a twist that a little simple arithmetical practice will solve.

(Ounces to grams? Celsius to Fahrenheit? Kilometers to miles? Do you shy clear of such annoying manipulations, or welcome the opportunity for proving you know how to solve the problem?)

Speaking of computers, many users are familiar with flow-charts, yes-no gates, and such tools of the trade. But even non-PC householders could borrow the idea of exploring "what if" possibilities in certain decision-making. Don't be afraid to put options down on paper, and work out what might happen "if" — perhaps with a rough indication of probabilities entering into the final tabulation. Sounds deadly — but may pay off.

How many people put their money in a bank, or a Money Market account, or savings bonds, or whatever, with no real idea as to whether they have chosen wisely? It's easy to say "I can't bother with such things" — but the difference in payoff might be surprising — and we are not talking only about financial returns. Such things as security, liquidity, short-term vs. long-term gains are all ingredients in the data mixture you use for decision-making.

There is one most important tool in problem-solving, whether the target be a cryptogram, or a word game, or a decision to pay the rent as opposed to take that trip. This may be termed "critical analysis."

In the sixties, my offspring berated me for "over-analyzing" and seemed to consider "critic" and all its derivatives as dirty words. They thought it somewhat immoral to say that A is nicer than B, or B is better-looking than C, or C is brighter than D. But without "criticism" what would the newspapers do for reviews of books and plays, or give help in deciding to attend (or to pass up) a concert or an art show, etc. Obviously one should see for oneself, if possible, and make one's own judgment accordingly — but such judgment is a form of criticism.

We can argue about the "scientific approach" to such things as problem solving, but a proper attitude toward such a pursuit probably entails a certain way of thinking,

specific ingrained habits, and even several personality traits or psychological outlook. One may have to learn or develop some skills — statistical analysis or facility with probabilties may have to be formally acquired, while a gut feeling for rightness or wrongness of an answer may be more casually picked up.

But there is no excuse for dismissing the fundamental rules and replacing them totally with hunches.

An analyst should be more of a skeptic than a bandwagon jumperoner, more of a practical realist than an idealistic dreamer, more of a questioner than a blind accepter, yet more confident than hesitant — which may seem to be an inconsistancy. Many of the giants in my field were admittedly egocentric — but I hate to admit that certain undesirable qualities are necessary. Even the show-offs I hint at were undeniably "lovable."

The questioning attitude, the insistance on considering the statistical data, the impatience with stupidity or error, the seemingly inhuman demand for the "right" answer — these are hardly the ideal nice-guy attributes. But it doesn't imply an overweening attitude, to be an honest searcher for "the truth". It either "reads" or it doesn't read — and I know of few fields where there is so little chance for bluff or bravado.

I remember the military boss (who knew very little about the technical field) coming in occasionally, and leaning over the shoulder of my mentor Dr. Kullback, invariably asking "Can you read it, Captain?" This innocent question still evokes what might be the clincher in many of the arguments met at the crossroads of decision — will it work, or won't it work? Is it right or is it wrong? In the final analysis — and the military man knew not whereof he spoke — "Yes, but can you read it, Captain?"

NOTES ON CHAPTER XIII

(1) One type of digital square (which could be used in Swagman) is easily made by writing down a random order (2 1 5 4 6 3, for example) and off-setting it for each subsequent row:

2 1 5 4 6 3
1 5 4 6 3 2
5 4 6 3 2 1
6 3 2 1 5 4
3 2 1 5 4 6

The above arrangement is called a Latin square and results in each column having no repeated digits, as well as the rows. In "real life" we found that in certain cases of usage some specific phase of cryptanalysis might be made harder by use of a Latin square, but there were also advantages and disadvantages to a non-Latin square (such as we usually find in Swagman.)

Without going into detail, we might say that certain recoveries were easier with a non-Latin square (which had certain columns in which repeated digits also implies the absence of certain digits) but that the "answers" had several ambiguous possibilities.

On the other hand, Latin squares might not offer an easy break-in, but when we sweated them out, the answers were unique. The British "swings and roundabouts" again.

Of course in Swagman, we have a kind of Latin square in that columns must be unique as well as rows, even though not created by offsets.

It may be rather tough to create truly random squares with this property, especially larger dimension squares, such as 10 x 10; it is much easier to offset a sequence than to scribble down "random" row after row (using each digit once on a row) without getting a crash in some column, and painting yourself into a corner. Try it on a 10 x 10.

(2) Did you spot the phenomenon? If not, prepare a chart which shows what digits follow any given number, as well as the frequencies in positions two away, three away, and even four away:

You will find the adjacent counts very rough, with certain relationships completely missing. For instance there are NO dinomes such as 11, 22, 33. And, for example, the digit 3 shows occurrences "next door" of every digit besides 0 and 6 (and of course 3.)

Moreover (and this is tricky to spot) while there are a lot of 01 there are no 10 occurrences, and in fact if you see a given dinome, you will never find its reversal. Plenty of 60, no 06.

A little thought should lead you to the explanation. In an attempt to suppress XX and X-X patterns which are so much help in Pollux solution, I carefully selected the next available alternative for dot, or dash, or X — running through all the dots in the order of 3 4 6 0, and so on.

It is interesting to compare the two versions, since there is one large chunk in common. If you stumble on the explanation (or more properly put, if you discover and explain the phenomenon of my second version) you can take advantage of the rhythm of variants, and read things easily.

I am tempted to submit this sometime to the CRYPTOGRAM, labeling it Morbit, to subsequently wonder how many hairs have been pulled out in frustration. But since it is a rather stupid trick for the constructor to pull — a veritable boomerang — I should reserve this for a "what not to do" example.

Chapter XIV

PROBLEMS TO SOLVE

THE exact type of problem will not be spelled out for you. A hint or so may be provided if it seems necessary.

No. 1. What phenomenon do you note in the following message which may aid you materially in solution — even though it is a simple problem to begin with? And what caused it?

JMWQ MW SBNGWZ VBU VA UMW XVKU KAUWO UMW ZWIKJKOW
YOVLLPQR JVBIZ EVB LKE JKLMPQRUVQ IWGU MWOW?

(Hint: There are a couple of proper names – usually the CRYPTOGRAM marks such with an *.)

No. 2. In cutting for deal at bridge, the other three at your table have drawn the Queen of Diamonds, the 9 of Spades and the 8 of Hearts. Is your chance of dealing slightly better than one in five, slightly worse than one in five, or exactly one in five? This is such a trivial problem in probability that I hesitated to include it, but it may remind you to remember the two worlds of "with replacement" and "without replacement".

No. 3. You are given a Caesar tip of FYR for a three-letter crib in a mono. What plain word did you get?

(Hint: I hate these three-letter Caesars – for more than one reason.)

No. 4. A CRYPTOGRAM offering by "BARBARA" with a Caesar crib of HTZUQJ:

IREYI SRBJU TIWRB XYERV HRTIJ UYDYH ZYWJM YXXTB
ODHYI YBJJU YURDD PWCKD EYMTE EJUTB GCZPC KMUYB
YLYHJ UYHYI RDCMY HCKJR OY

No. 5.

NSWIJ NPDIN ZNWJK ZGMYI MYNJY ISJSH DPGCY
IOLJN NXNUA HVOEU ICUCJ ZDGEF BTRKM NFPLN
IXMQS HKRDR LNEVQ UPRMY EULPY CQWNG QU

(Hint: Would you be surprised if this turned out to be a digraphic system of some kind? You will find two specific letters have a certain rhythm. Try SQUARES as a beginning, and find FREQUENCY someplace.)

No. 6. There may be no sensible cryptographic explanation, but is there any describable mystery about the following sequence, apart from the fact that each letter is used once?

F A B U K P V G H C L Q M W D R X N I O Y J Z E S T

(Hint: Don't struggle with this – it is included just to see if you react to unusual rhythms. Nothing to "solve.")

No. 7. Are you familiar with the classic anagrams ROAST MULES and CART HORSE? If you don't know the answer, the first one is hard.

"A RUG'S NOT A sort of this" might be part of a clue from a *The Nation* puzzle. If you got an answer, there is a 50-50 chance you might have trouble with fitting it in. Why?

No. 8. Can you see anything about the following that might lead you to believe you have some of the "enemy's" crypt components?

A D J B E R Z U V Q O Y P K F U G B Z C N V I Q A J
D L W R E M X H S T W E R O B S F D K Q G A T U P I
C Z J V H X M Y N L Z B G C I Z V X W R

No. 9. Can you recover the keyword of the following sequence?

D I S E B L U K A J T O G P Z R H Q W F N X Y C M V

No. 10. Two messages show unusual characteristics. What kind of a system do we probably have?

#1 96447 31817 02510 29822 04513
#2 96449 46617 02725 06176 81007

No. 11.

Message:
32684 32963 56523 66429 53674 65361 29417 13366
64405 84067 43493 14162 56694 93173 52434 93240
42404 86933 36436 13637 44505 95268 33445 04861
57374 34663 72575 36349 63575 56954 46734 66856
32264 25046 44583 17351 32363 25164 60366 85820
74466 75632 27634 64339 42297 46746 59525 04872
56435 84863 71453 76151

(Hint: Certain numbers, such as 9 and 0, seem less frequent than others, as you have undoubtedly observed. They might be a good place to start looking for repeats – three or four digits, at least. The fact that I suggest starting with less common digits is merely because a quick eye-search may be easier. You will probably find more repeats involving the common digits, but you will naturally expect quite a few short repeats at random. Those who know Nihilist Substitution should not be fooled. Close, but no cigar.)

No. 12.

CYEBL SEDTG IELEB EEXOC REMYT REOEH NLMEE ALSET
EARCG PTENI GNASR INYIS CTMTH WIBEV SSYSC AGAOO E

(Hint: With such a "favorable" length, try the simplest thing first, and hope for the best.)

No. 13.

STLIF HIENT ETSEH IERAN NSSSL TETHT PEEOE NTRIP
SAILT TOAES ILONF CRIAE TIOIA PTTNU HEIRO SRENO

(Hint: Look for -POSITION-. Also, for what it's worth, you would be unlikely to find this specific category of encipherment having the same letter count as No. 12. A word count of 49 would be equally unlikely.)

No. 14.

1. CTKCITINRHCX NIT IQYWCFIT PIQYX LCTIHHA WRHHCTP
2. OXYZYP WVGGJONW FYNAP FXYBXOW' GXYIX YZOXXHXPN
3. QZWHI VIQR-UARWIQ PHQIVRJF, QGGVHF ZHQI FPQIVRJ XQWQ
4. GCDMOXZNO XG AMZWKXI ADK OIXEZ ODMGC XAKEYXZ PIXYT
5. YGRMGPT KPFOHF DYKRJMOGH GH JOYT

(Hint: Courtesy Paul Derthick. Don't worry if you get a garble. These things happen in real life.)

No. 15.

5 24 - 34 10 51 25 28 2 3 6 50 26 14 - 13 - 51 23 93 31 8
16 27 65 48 9 - 46 7 74 44 27 - 1 19 64 11 - 12 93 12 18
20 21 - 22 57 15 - 41 95 69 - 102 - 40 134 30 - 117 35 81
- 38 77 33 - 11 59 29 53 37 56 78 88 - 89 103 - 58 61
42 - 105 45 77 7 66 - 49 101 68 60 - 93 35 36 - 94 62 51
80 133 55 84 - 63 76 107 90 71 133 82 6 - 72 85 135 75 -
73 112 69 - 104 105 3 34 - 17 91 92 - 96 76 99 - 70 3
93 - 12 90 87 32 - 74 125 - 28 23 102 51 1 5 51 109 - 25
8 - 63 59 106 108 111 - 67 116 - 3 135 83 109 111 56 18
89 78 106 105 - 37 91 - 102 28 31 45 110 111 - 107 103-10
135 49 6 13 - 48 98 43 2 95 54 99 90 - 48 96 - 97 2 3 -15
110 115 - 22 4 - 100 76 20 - 122 82 93 - 60 59 - 83 84
126 133 56 34 33 - 39 80 81 - 112 108 98 - 28 80 11 12 25
67 108 10 - 123 15 49 90 10 84 7 - 130 49 16 83 - 79 22 6
3 - 86 82 12 11 107 27 116 105 - 63 87 46 98 34 - 79 48
70 99 - 69 78 29 53 1 25 22 44 - 3 48 69 74 71 6 - 54 53
8 - 29 30 24 27 10 6 - 70 55 93 - 63 78 106 108 34 - 120
110 51 6 20 13 12 33 - 1 61 71 - 34 5 17 39 25 51 53 29 8
93

(Hint: This is a Pseudo Book Cipher. KEY PAGE SHOULD BE NUMBERED CONSECUTIVELY 1 TO 137. You will find the last part of the key page accommodates a normal alphabet, in case the key chosen does not contain a certain letter. And it doesn't cost much to check 1 2 3 (the first numbers of the key sheet, of course) to see if it starts THE. Maybe you'll hit pay dirt.)

No. 16. Suppose in a large body of traffic you had the following frequency counts for the first several positions, dropping the last digit in order to make the numbers more manageable here. Admittedly the problem is highly idealized, and in actual fact you would have a much flatter distribution except for one important fact that a fussy statistician might poohpooh. (I could have cooked up a different presentation and not worried whether I was stretching the statistical approach or not, but this is more fun.)

First letters of messages:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
	67	64	49	70	58	65	71	90	56	82	66	87	71
	N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	63	69	83	77	88	53	49	72	74	85	79	80	67
Second letters:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
	69	70	81	52	45	66	62	71	75	69	67	83	78
	N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	66	49	77	72	73	62	68	72	82	77	77	55	64
Third letters:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
	83	76	50	71	82	59	69	74	65	74	84	85	70
	N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	48	84	51	79	80	72	75	78	78	85	79	68	79
Fourth letters:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
	58	76	69	80	76	53	70	80	84	81	69	76	83
	N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	85	78	73	81	53	49	78	86	81	82	74	71	70
Fifth letters:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
	57	66	81	81	49	69	73	63	50	72	80	74	67
	N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	59	68	80	88	56	59	54	64	63	73	83	85	76
Sixth letter:	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**	**J**	**K**	**L**	**M**
	77	84	48	54	83	83	79	72	69	89	88	83	87
	N	**O**	**P**	**Q**	**R**	**S**	**T**	**U**	**V**	**W**	**X**	**Y**	**Z**
	76	67	68	86	80	71	50	59	69	75	72	72	83

Could you hazard a guess as to what you might expect to find in the seventh letter distribution?

No. 17. Just to prove you can solve monos in foreign languages, even though you think you don't know the language well enough, here's one in German. If you think of

the several forms of "the" you will get more help than we do from our constant "the" in English. This one is by "Le Crapaud" — and a crib of RHBG (in Caesar form) is handed to you on a platter.

QXR GMBCPWBXTD; GMB QRS KFCLSSRAOLF CMUUDR SLA
CXPW SXD QRA RXAKRUARA DRXURA XA QRB
FORBCXPWDCKRXPWAFAV GRBDBLFD SLPWRA.

No. 18. Usually amateurs are told what type of cryptographic problem they are faced with — and as a result, I sometimes argue, if I happen to be a guest speaker or something, that people are missing a lot of diagnostic fun. Many, if told that this is a Bifid, or a Quagmire, or whatever, know exactly what to do to solve the message. But a lot of this group won't move a brain muscle if the label is torn off.

The other thing I argue about, is whether or not even a not-too-experienced solver should not be able to develop an approach to a problem, once the method of encipherment is explained, and an example of the cipher is given for solution.

Some, instead of asking themselves "what is the problem, and how do I go about untangling it" expect to be told exactly what to do. "Take such and such a count, and if you find thus and so, take your pencil and write the message out in this specified manner."

Now, for those who want a little "don't tell me" exercise, the following three messages are given. There is ample evidence in each one for accurate diagnosis, and with a certain amount of effort, you should be able to effect a complete solution, even if you aren't sure even what to call the problem.

A. CLOYN OTEFT EIUCF IOMNI ECTSM PTEHS HALHE NFEAE MIIDF
LAOON EUDLN REDEE IDTTT EHANE VNGEA TNMAR TAEAC TAHVL
CENEO TOXNR

(Hint: A Caesar tip of GXGTA is provided. I will give an extra hint by saying the message starts out by talking about a French ruler.)

B. MAJTA VGZTZ MUMAJ QTHNS UMNTH SSZVC AYVAH CSJTF FVUTH
FPVXH DAVUU TXMED HWNVF AVBVF FVPVC BVQMN CDDXF TWTAV

(Hint: A Caesar tip of FSDTSJ.)

C. VZWRS PKKNQ JRKQL MLWWL ITHDA YGZCG AJMCJ RKFMM ACAOJ
EUGRJ ZTKAB HAYGX OPYNA ZWQDJ LETKJ TOLXJ EYBZQ QLPDK
NAIGR EJPKA OLXWT GITNP AJAZJ RKYKK RRDKP RHDAK JRKCA
CFCWB NLTBL DBOJE ZHAZW VEQMP LXVOJ EBAAC EZZWM BLKCD

(Hint: Do you notice any repeats? This is talking about Alan Turing and his ACE computer.)

In the above problems, your main task is diagnosis. If you can read them, fine — but you may want to peek at the notes on this chapter and get a few more hints (unless you are an accomplished solver.)

No. 19. Here's a problem that many solvers have no trouble with, when told what it is — but the phenomena of which might escape such solvers, if they were merely given a text like this:

```
TTWMT QABWT WYTIA BDTQC NOKBG TQWKY NXNEX KIRQH HOUBN
RFVNY BIGKW PDRFN DOXQO IHELE BKHRM TANFS TZICF
```

Note the TW repeat and the AB repeat close to the beginning. It looks interesting in that they both agree on a seeming cycle of 8, but if you pursued this course, you would not be happy with the results. No QUAGMIRE, or whatever you choose to call such a polyalphabetic system.

What things have you noticed that might get you on the right track?

No. 20. Since it is in numbers, I could have written the next problem out in 5-wide groups, which would have been an extra challenge to the diagnostician.

However, I give it in the dinome form originally presented, and to make up for your not knowing exactly what type of problem it is, I will give one beginning word — FOLLOWING. The phenomena (because of the crib) should give things away.

```
F  O  L  L  O  W  I  N  G
40 13 61 93 48 31 91 63 06 90 25 49 97 23 35 03 38 59 18 43 64 12
28 96 05 51 11 91 10 81 69 08 47 33 59 08 45 93 07 83 54 76 13 46
04 87 18 95 39 75 25 26 67 43 20 35 61 19 65 48 12 41 16 51 63 86
62 83 25 27 09 18 12 57 03 64 95 28 74 96 29 21 35 95 69 76 96 17
55 87 23 64 70 99 01 51 36 74 68 43 00 76 55 99 69 42 04 277 5 48
16 09 52 26 13 77 89 58 18 68 94 83 45 10 41 59 67 93 84 61 77 63
76 45 22 17 83 91 38 47 13 08 79 51 12 69 28 48 10 64 96 44 35 18
76 07 39 00 69 13 99 60
```

No. 21. Although nothing so complex as the ADFGVX system mentioned, here is an opportunity for "special solution" — the main entry into the WW-I system which combined substitution (by fractionation) and transposition.

A general solution was not developed until many years later, but the phenomena you see in the following much easier example, enabled occasional reading of messages. Can you spot the shortcut?

Solving it by "standard" procedures which ignore the funny business deserves only half credit. Message sent to one Ambassador:

```
IKEWR CDERF OADML NEFNT HMYIN ATABA SGTHA
ATAIL INUIZ BOISN IEDAR SUEE
```

Message sent at the same time to another Ambassador:

ERFOE SMLIK EWBUD SGTHN ATNEF NCAMY ZBOIH
DIEDA RSUTE INATA NAAIL IARI

No. 22. What do you make of the following four messages picked up on several days watching of an agent network?

A: NBDRE XACFS OYERH XWEVC EKINX THXYE POOXL
UNURT OXXOF NADRI EXOAT SLPUE EXXS

B: NVERE XXXRE ACILF ONRGO HMTTX ACXYX NOTDO
XANXA YFSSA XPORX XPXWL OXSER AEDS

C: NSXRW XQXEN RUIXY ALONU ERIXX FIRNT EREOC
DXXFX QOYUE ESHRX ERXIY VXAXE INDS

D: NDRRI XNMAT IXEEU ATNNI XCXCA ONETL AICYS
EXONX ASXNX DXXIX EXMXR CXMEO TEMU

No. 23. What do you make of this? (I'll give you a tip of INCOMMON)

SEASS DOEWA LSAWR BSBEO RDOEW LESEO SAWDB RASEO
DSDDS LSDDA SLWLR EBROE WLDER ASEDW DDRBS ALSDL
SLDWW DLSWL SESAS DRBDB BDBRA SSESA WDDDS LDASL
OLWAW DDDLS WLLSE SAWAW DWSED LLDSA RBOED BERSE
WEODE SDWLO

Hint: In taking a running frequency when I do not know whether the message could be digraphic or not, I sometimes write down a normal alphabet, and fill in the sequential letters alternately on top and bottom – thus above S, E is written; below E, A is written; above A, S is written – thus the top distribution is of digraphs on the cut, and the bottom distribution shows digraphs across the cut of two.

Perhaps the distributions will match (monoalphabet?) but perhaps the count showing digraphs will be quite different from the one crossing the border line. It doesn't cost anything to get the extra information.

No. 24. Simple probabilities to keep you aware of such. At dice, your point is 9. What are the chances of throwing 9 (you win) as opposed to throwing 7 (you lose)?

No. 25. What do you make of the following? (I mainly ignored this type of problem in the previous chapters, but it's a good chance for diagnostic discussion.)

CKZGK BUYVO OIBPZ DBJBO CVLEM FJBPU ICNRU
ODECG XOCTL LDSDD LYRVS LLFYM WAY

Hint: Occasionally you may notice peculiarities in the Delta stream (the running difference between consecutive cipher letters – with a known or assumed alphabet, of course) when you can't seem to find much in the cipher itself.

Such a phenomenon can also be spotted as an isomorph, if you compare one stretch with another. And you may be lucky enough to have several such combined —including roughness or repeats of cipher.

No. 26. I added this almost as an afterthought, since I really don't expect many of you to attempt to recover the message, so-called. But you should easily come to some conclusion as to what gives, and may have some fun venturing a couple of guesses about text.

```
GKIUA VBEFW FCXQH YACZD JBVCE WEMRX LYABZ AVFCB WHXEF
EUGEW RHXAD UZZUD JDULX CIBXC FYFBC YBKZA EHUAS XCEOU
AUGKI WAVGB GXDYH LGZFJ ZDFPX CYAJN XBAZP OVCIW DFOVD
ZAQJU CKYIC DUAJX ESJUE TWCMG VEAWG RFVDH WICDV APWFH
QXQYG QLZCG ZHTDY MXCEP ZDUIA DVEJX BAYAZ CIAYC FWHLI
VAXBP OUDJV HRLWD GXCMY GBZCB IXFLZ GKIYA VDMJU EKWAJ
TUBKX GSDYB ZGJIY EKUEI JVCAX PAIVC LZHGR WFBZF LUPVF
BKXBL WEALY BKWEA LZAKX FAKXB L
```

Chapter XV

NOTES ON PROBLEMS

LET'S SEE how you did on the problems in Chapter XIV. If you quickly answered them all, there is obviously no point to your reading these notes, and you can consider yourself a qualified cryppie and Problem Solver.

No. 1: The most striking phenonemon consists of the long word in alphabetical order (with a few omissions) JKLMPQRUVQ.

The obvious reason for such an occurrence is that the monoalphabetic substitution is effected by a slide of a mixed keyword alphabet against itself — the keyword of the plain sequence starting over J in the cipher strip.

Several assists are given by this — such as knowing that the keyword has N, O, S and T — and has a repeated letter in sixth and last position.

In otherwise difficult cases, this phenomenon may provide a clue to certain letters; note the frequency for the stretch before J in the normal order:

B	C	D	E	F	G	H	I
4		2		2		3	

It is unlikely that cipher I is plain Z, so may be in the keyword. H or F could be under Z, with G or E as Y, and C, D or F (and possibly E) for plain X. All these possibilities are worth juggling around — and THE should be a three-letter group starting with one of the JKLMPQRUV set — making UMW look pretty good.

No. 2: I'm sure you didn't fall for the "without replacement" booby-trap, but some people might work hard counting the two higher Queens (by suit) and the eight Kings and Aces, and say "10 out of 52?" Of course you said, "10 out of 49," which is slightly better than one chance in five.

No. 3: Cheer up — we'll be getting real problems to solve soon. But just to throw in a warning to be careful, and not ALWAYS jump to conclusions, I put in a CAESAR that has two answers — HAT and BUN. You may have missed the latter because in extending, most people keep progressing normally — FGH... etc, then a few from YZA, and a few from RST...

You might say to yourself, "why run down the entire 26 letters in the first position, then the second, and finally the third, when you might see the answer by writing only part of the extension under each letter, hoping to see something."

Fine. Try the quickest approach, but be prepared to back-track if you become suspicious that someone may have "had" you.

No. 4: Quick inspection shows that this is a so-called Patristocrat — it is obviously monoalphabetic (note there are 17 Ys out of a limited text) with one three-letter word occurring three times. The way the repeats cross the separators shows that this could

not be an Aristocrat (with word length) even apart from the "all-five-letter" words (apart from a dangling 2-letter word.)

You undoubtedly recovered the crib (or tip) as being COUPLE. If you assumed that the repeat was THE, there weren't too many likely spots to place the crib, with the W and K ciphers satisfactorily low, to be C and U. The D and E for P and L are not unreasonable, although the C for O is a bit low. So you had a Plain-Cipher set-up of:

```
PL: ABCDEFGHIJKLMNOPQRSTUVWXYZ
     W Y  U   E  CD   JK
```

which is almost too good to be true — the area of keyword (and the non-keyword remainder) being almost exactly pointed out.

And you couldn't possibly have missed filling in THE———COUPLE, even without the PP.

No. 5: Playfair, despite the presence of a few J's. You will note that I and J alternate through the message — perhaps the maker went over the text and changed every other I to a J. The rhythm tells you 25 letters, not 26. Not a cast-iron certainty, but would you expect 11 heads in a row in so few tosses — or more suitably, alternating heads and tails?

No. 6 isn't a message, but was put in to see whether you could quickly spot the rhythm of the five cycles of alphabetic order. A, F, K, P and U all start a left-to-right sweep in alphabetic order. It could be that someone thought this was a good way to get a mixed sequence.

No. 7 consists of a couple of old saws — the anagrams for SOMERSAULT and ORCHESTRA. I put in ARGONAUTS and ANGOSTURA because my lunch gang bit on it once, and struggled with the "obvious" but "wrong" entry.

No. 8 shows that phenomena can begin any place in whatever you are looking at — not necessarily the beginning. If I had used 26 unique letters for the first stretch you would have spotted it, but because there are two "indicator groups" (of the A-J and Q-Z range) the sequence of "no repeated letters within a span of 26" starts on the 11th position, with two sequences and a couple of final indicators. Of course a frequency count would have told you something was up, since it is so flat as to be remarkable.

No. 9: I trust you proceeded by seeing what happens when you assume some sort of box with a transposition key — trying J next to K (not very exciting) and P next to Q (much better, giving GH) etc.

The 7-wide keyword (appropriately KEYWORD, although it is too short to legitimately recover) gives the numerical transposition key of 3 2 7 6 4 5 1.

No. 10: One can hardly escape the fact that there are 4-digit hits between the two messages, placed at the correct cut for a 4-digit code. The chance of this being so is increased by the fact that the differences beginning at positions 5 and 13 (or at the second and fourth position of 4-digit code groups) are both 8952 — using non-carrying arithmetic, of course. Had one of the differences been 2158, it might have taken a trifle longer to observe the switch. For this reason, once a specific and regular size of the

groups is suspected, differences are usually taken in a manner designed to obtain the lower value of the two possible ways of subtracting — top from bottom or bottom from top. The so-called minor differences (naturally on the beat of 4) would be 0000 2158 0000 2158 3756 2351.

No. 11: The most cursory examination of the cipher shows a decided bias in favor of the numbers in the area between 3 and 6. And one sees repeats all over the place — for instance 4931 — separated by 8 digits. Jotting down the positions of just a few of the hits, I note:

4931	positions 53	and 65	Interval	12
4450	101	113	between	12
6685	156	194	positions	38
5048	115	235		120

Note that three out of the four occurences are compatible with the interval 12, and when one looks at the three-digit hits, one builds up confidence in such an interval. So let's write the message out 12 wide:

```
3 2 6 8 4 3 2 9 6 3 5 6
5 2 3 6 6 4 2 9 5 3 6 7
4 6 5 3 6 1 2 9 4 1 7 1
3 3 6 6 6 4 4 0 5 8 4 0
6 7 4 3 4 9 3 1 4 1 6 2
5 6 6 9 4 9 3 1 7 3 5 2
4 3 4 9 3 2 4 0 4 2 4 0
4 8 6 9 3 3 3 6 4 3 6 1
3 6 3 7 4 4 5 0 5 9 5 2
6 8 3 3 4 4 5 0 4 8 6 1  etc.
```

Even without completing the write-out, one notes that many of the columns are severely limited in their composition. Listing them:

Col. 1	Only numbers	3 4 5 6
2		2 3 6 7 8
3		3 4 5 6
4		3 6 7 8 9
5		3 4 6
6		1 2 3 4 9
7		2 3 4 5
8		0 1 6 9
9		4 5 7
10		1 2 3 8 9
11		4 5 6 7
12		0 1 2 6 7

If you'll pardon the expression, all columns are funny, but some are funnier than others. You will note that many of them have a maximum of four digits AND THESE FOUR DIGITS ARE CONSECUTIVE. Which columns? 1, 3, 5, 7, 9 and 11.

What does this suggest? Four rows of a rectangle for digraphic substitution? (A 12-wide additiive would obviously have to have been added, in effect, sliding the 1-to-4 coordinates to positions such as 4-to-7.)

If so, one needs only seven columns (giving 28 cells) to accommodate the alphabet — with perhaps a couple left for punctuation or something.

This would account for the wider spread in the numbers of the second position of digraphs. And sure enough, the consecutive numbers can always be encompassed going around the corner at times, such as the 8 9 (0) 1 2 of column 10 or the 2 3 (4 5) 6 7 8 of column 2.

It is a relatively simple matter to discover what the key must have been to form such a cipher — and lo and behold. The key behaves in the same manner as the cipher — implying it may be literal key put through the same matrix as that used to encipher the plain text.

Readers who took my suggestion about checking on Mr. Friedman's *Elements of Cryptanalysis* might refer to Paragraph 57 in that work, which describes a system used by the Hindus in 1915.

I took the liberty of adding a word separator, which makes the problem even more simple.

Also, those who know the Nihilist Substitution may have been fooled for a few minutes, but would have sooner or later come to the conclusion that, while the first position of a digraph is severely limited, the second position is less so. Thus, while the Nihilist wouldn't suffice for an answer, it should suggest the general "shape" of the problem.

No. 12 is obviously a transposition. Without a crib, such a problem might be difficult, if the word length were 85, say. But with 81, you can hope that you are dealing with a completely filled 9 by 9 box, and all you have to do is try to line up adjacent columns to show good digraph values.

If you wrote the cipher vertically, nine letters at a time, you may have "seen" the answer in the top row, which contained the letters CGOOLPRTY, clinched by the fact that there is a convenient Y underneath the C.

No. 13 gives itself away by the fact that you see the letters of POSITION (in the right order, but with certain intervening letters) not once, but twice. It must almost certainly be a 9 x 9 grill. Remember that with odd widths — 7, 9, 11, etc. you must have a blank in the middle of the square in order to accommodate one quarter of the total letters at each turn of the grill. Thus the actual word lengths would be 48, 80 and 120 for the widths mentioned. The first box starts out with THE, as you undoubtedly discovered.

No. 14, as you could tell by the thanks to Paul Derthick, is one of his headline puzzles. The beginning words for the five messages are:

INDIANAPOLIS REAGAN AFTER THOUSANDS DOCTORS

The first guess I made was the S in No. 2, with SUPPORTS providing enough for solution. (The REAGAN pattern would have been an entry point as well.)

No. 15. You probably started out by assigning the value of A to the 13th position of your key sheet, since the one-letter word is A (unless it is I). The occurrences of 1 2 and 3 throughout the message look good for starting the "key" page with THE — and I will merely confirm that the key sheet reads:

```
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16   etc.
T H E F I R S T M  E  S  S  A  G  E  O F T H E  etc.
```

Therefore our problem message starts out with a two-letter word beginning with I since we will call all 5's I, all 6's R, etc. IT? IN? IF?

No. 16. You can pretend you are working on Zendia's new machine which they grabbed when some of the Nazis fled to their island.

You no doubt spotted the unusually low frequencies of certain letters in the first several positions (which would taper off to flatness if I gave you counts in positions further on.)

If you took a frequency count of the values by position, you would end up with a "kind-of" bell-shaped curve running, say, from about 60 to 90 — with nothing very exciting about that area, except its flatness. However, there seems to be a wide "handle" sticking out to the left of the curve, in the area of the 40s and 50s.

Is there anything interesting about these few letters that seem to be at the extreme low end of the curve, in one position or the other?

Well, for one thing, they are mostly pretty good plain text letters with good values such as E, A and O being present. Why are they so low?

Let's just list, say, the four lowest letters, by frequency, for each column:

Col	1	2	3	4	5	6
	C 49	E 45	N 40	S 49	E 49	C 48
	T 49	O 49	C 50	F 53	I 50	T 50
	S 53	D 52	P 51	R 53	R 56	D 54
	E 58	Y 55	F 53	A 58	A 57	U 59

Probably an exaggerated example of a few very stereotyped beginnings using a "reflecting type" device such as ENIGMA, wherein A can't be A (unless special precautions are taken.)

Now I'll admit I cooked this up, and had not access to gallons of traffic or a computer program to generate and count a complicated cipher from a machine — and the sample

is badly enough distorted to drive my fussy mathematical friends crazy. But the point is to make a point.

IF a reflecting device CONSTANTLY enciphered STRENGST GEHEIM at the beginning of EVERY message, there would obviously be no cipher S in the first position no matter how much traffic you counted. (Presumably the rest of the first-position letters would be distributed randomly on your count.)

Now suppose there was another common beginning — just GEHEIM.

There would be some Gs caused by enciphering STRENGST and some Ss caused by GEHEIM — but wouldn't S and G be statistically a bit on the low side?

If there was a third common beginning, things would tend to get sloppy, but since three letters were in the category of one-time impossible and two-time random there might be a shadow of the cribs to hope for.

Of course, nothing could possibly end up as pronounced as the cooked-up deal, but that was one way of giving you the problem of picking out three words or phrases from such a set of letters— as one has to do in solving certain puzzles that occur in crossword magazines under such titles as "Quoteboxes," which I described in Part I.

If you still haven't seen the words, start thinking in English instead of German.

No. 17. After you've done a couple of monos in German, you may find them easier than English. Der, die, dem, den and such are spotted a mile away, and the IE EI digraphs are sometimes spotted, along with the almost invariable CH (which is almost as much of a crutch as QU, but much more common.)

Note the length of words, which may be a giveaway when you are faced with a message in an unknown foreign language.

We didn't say that this was a K2 alphabet usage, but I'm sure you recovered it, if you got as far as UBERSICHTSZEICHNUNG.

No. 18. Simple inspection of the A problem almost guarantees the fact that we have a transposition of some kind. The 100 letters might suggest a grille, or a completely filled rectangle. But if you worked around the two V's and the single Y for a time, you wouldn't get very far with the rectangle, and would have limited success with the grille idea.

While you can see the natural progression of E V E R Y in several combinations at the end of the message, going to the Y in the first group, you would not get anything sensible in trying the 180 degree reversal of a grille.

But do you remember the Swagman, described in Chapter XIII? Even if you didn't know that system, there are plenty of phenomena to discover the method of encipherment, especially if you follow the suggestion about Louis, and come up with some "teenth"— in this case "four."

Being told it's a SWAGMAN makes it a simple matter for the solver. DIAGNOSING the system, knowing nothing about it, is a different matter. Can you see what the

real-life cryppie (at least in the old old days) might have been faced with, and can you sense what satisfaction is obtained by such a relatively simple analytic job?

Unfortunately (or fortunately?) the days are probably long gone when such a simple challenge as this might be offered to a professional.

Problem B is entirely different, since one quickly determines that it cannot be a transposition (with 13 Vs.) and is probably a mono of some time — a Patristocrat, as ACA calls it, because there is no word divider.

How to tackle it? The crib, and we are always grateful for such, has a repeat of N, so one merely goes through the text looking for a pattern of X--X. There are a few, but the first part of the message starts out that way, and good frequency, with the N being represented by A eight times and the E enciphered by V 13 times.

Considering the possibilities of various alphabets, if one writes a normal order, and puts the cipher-to-plain both going to and coming from the normal, one gets:

```
M       V                   A T                 J
A B C D E F G H I J K L M N O P Q R S T U V W X Y Z
N                   Y     A           O   E
```

If it is a K-1 type, with cipher normal, the keyword is in the L to V area; if it is K-2, which looks much better because of the JM sequence at Y to A, the (quite reasonably) has AT in it. Putting in the values:

```
MAJTAVGZTZMUMAJQTHNSUMNTHSSZVCAYVAHCSJTFFVUTH
ANYONE  O A ANY O    A O    E N EN    O  E

FPVXHDAVUUTXMEDHWNVFAVBVFFVPVCBVQMNCDDXFTWTAV
  E   NE  O A     E NE      E  E A      O O E
```

Perhaps it's asking too much, but the temptation of trying "Anyone who has (or had) any" is worth succumbing to, at least to the point of checking frequencies, or trying out the values. (Note the checking repeat of H in WHO HAS.)

```
MAJTAVGZTZMUMAJQTHNSUMNTHSSZVCAYVAHCSJTFFVUTH
ANYONEWHOHASANY O   SA O   HE N EN    O  ES

FPVXHDAVUUTXMEDHWNVFAVBVFFVPVCBVQMNCDDXFTWTAV
  E   NE  O A     E NE      E  E A      O O E
```

I quit the above operation when I saw the "about the" — and I'm sure you solved the alphabets, too.

In Problem C, I suggested looking for repeats — and what a windfall. You probably marked four cases of JRK, at positions 11, 35, 115 and 131. There were lots of other 3-letter repeats — HDA, AYG, OJE (3 times), AZW, and OLX — but the lucky presence of so many JRK's should clinch matters.

If you checked the intervals given by the JRK repeats, you would have 24, 104, 120, 80 and 16, wouldn't you — subtracting 11 from 35, etc., all the way up to 115 from 131.

Is there any doubt about the answer being an interval of 8? Let's write it out at that width:

```
VZWRSPKK NQJRKQLM LWWLITHD AYGZCGAJ MCJRKFMM
ACAOJEUG RJZTKABH AYGXOPYN AZWQDJLE TKJTOLXJ
EYBZQQLP DKNAIGRE JPKAOLXW TGITNPAJ AZJRKYKK
RRDKPRHD AKJRKCAC FCWBNLTB LDBOJEZH AZWVEQMP
LXVOJEBA ACEZZWMB LKCD
```

Notice the JRK's line up (naturally) as do many of the other hits. Almost certainly we have a polyalphabetic substitution on the width of 8 — what some call a QUAGMIRE, although as yet we do not know what type (in effect, we do not know how the alphabets were prepared and juxtaposed.) But trying type III, which uses a mixed plain sequence against the same cipher sequence, you hit pay dirt.

No. 19. An experienced cryppie might try the right approach on the basis of familiarity with the Bifid system, as detailed in Chapter VI.

There are several road signs. First of all, the absence of J, in 95 letters, while not conclusive, smacks of a 25-letter alphabet, which one might get in a 5 x 5 square — but several systems use such.

However, while the AB repeat is on the digraphic cut, the TW is not, which may or may not be critical. When split repeats such as the TW...AB..TW...AB occur, they are very helpful in Bifids.

The reason for this is as follows: In an odd width, such as 7 or 9, one often gets repeats such as

```
|. . T W .|. T W . .|     or     |T W . . .|. . T W .|
| . A B . | A B . . |            | A B . . | . . A B |
```

where a pair of letters is repeated which involves a letter from the bottom row and either a letter to the left or to the right on the top row — such as AT or TA above.

In the first of the above possibilities, the plain values over T MUST be the same, and the letters over the A MIGHT be the same.

In the second case, the plain over the A MUST be a repeat, and the value over the A has a chance of being the a repeat.

Suppose T has the value of 12 and A has the value of 35. In the first example you will have a repeated plain of 15 and in the second possibility, the guaranteed repeat would be 21. If this isn't perfectly clear, doodle around with your pencil.

Obviously, in good English plain text, the letter E might occur many times, and since there are only five letters which have the same row name (including E) for the top row, and only five letters which have the same column value, for the bottom row, there are

25 combinations which will yield E in such a position, instead of 625 combinations (any letter on top and any letter on bottom.

By the same token, if an E is in the "other" position of some cipher value having the same column numer as E's row name, there are only five ciphers possible on the bottom, having E's column number as their row name — another 25 out of 525.

This is all much more simple than any description can make it. In any event, if we examine the text, and try to juggle the cipher hits so that they fall together, a width of 9 seems the only possibility:

```
T T W M T|T W Y T I|
         |         |
Q A B W | A B D T |
```

When we search the text, written out on the 9-wide cycle according to Bifid rules, we find there are three more pairs repeating from bottom to the top right (GN, OR and BH) and four pairs reading from the bottom to the left, or preceding value on top (TT, QB, NK and HN).

This makes a rather impressive 10 hits out of 95 comparisons (9 for each 9-wide "box" since we must also count the last letter on the top row going to the first letter in that box.)

(Play around with that pencil again, and you'll see the T....I on the top of the second box would give the same result as any position where I on the bottom leads to a T on top. The plain value would have the row position of T and the column position of I in either case.) Now we have rather conclusive proof that we have a Bifid — and what do we do with it?

Note that our text MUST begin with T, since the first and last position of the first box are T — according to the reasoning above. Of course the text could start THE, with a two- or three-letter repeat of plain right after, such as THE RElics aRE —.

But another possibility is that a seven-letter word or phrase is repeated, such as "I'm going, I'm going." — or any word of the BANG BANG type.

It would be an easy thing to check, so let's assume a repeat, and start the 8th position with the same as the first, the 9th and 2nd agreeing, etc. (Beginning with 11th, we MUST have a repeat of the 4th onwards.) We'll arbitrarily call T 12. But notice that to start our repeat at the right place, we would have to change that 2 to a 1 — making T 11.

Also, if our 7-letter repeat starts with T, W (as the last cipher in the first box of 9) has to start with a 1 also, and we can safely call W 12. So why not call M 31 at the same time? And since the final W gives a plain text above it of W, the seven letter word must start TW, making Q on row 2.

```
TW      TW          T
T T W M T|T W Y T I|
111112311|11123 11  |
 Q A B W | A B D T |
12      12          11
```

At this stage, one could work with the entire message and sweat out the text, but the seven-letter repeat must be something that should ring a bell.

I looked up the TW—words in one of my lists, and found such as TWIDDLE, TWIGGED, TWINGES, TWINING, TWINKLE — Oh, OH.

And the lonesome T is in the right place for the next word.

The above "solution" sounds contrived, but most of the analytical statements were sound, and the seeming flights we took could have been back-tracked if we got into trouble.

20. Experienced solvers would have suspected a so-named Homophonic substitution, and made short work of it. But the alert solver who has never heard of such a beast could still find something "phenomenal."

Let me put the dinomes of the crib in order of numerical ascension along with the plain value:

```
06 13 31 40 48 61 63 91 93
 G  O  W  F  O  L  N  I  L
```

The L N order should be the one to catch your eye, but I'll assist you by putting in a few slashes to mark off seeming boundaries. I also append the values in the normal alphabet, although I could have improved upon this.

```
06 13/ 31 40 48/ 61 63/ 91 93
 G  O/  W  F  O/  L  N/  I  L
 7 15  23  6 15  12 13   9 12
```

You surely have noticed the almost exact relationship of interval between the key numbers and the ordered letter in the alphabet, within the four blocks delineated.

Four blocks in 100 dinomes? Doesn't that suggest four sets of 25? And doesn't it explain the few "misses by one?" No J.

Change all numbers to modulo 25 arithmetic, and you can even go around corners as in W F O = 22 (31 39) under 31 40 48.

Putting in all four normal alphabets, starting at a different place within each 25-wide stretch, is a simple deciphering chore — all on the basis of a one-word crib.

No.21. I'm sure you noticed the many short repeats between the two messages, seemingly involving a certain rhythm. Let's write the text out and underline repeats — some of which, of course, may have an accidentally "hitting" letter to make the rhythm more confusing.

No. 1

```
  A       B       C       D

IKEWRCDERFOADMLNEFNTHMYINATABA

  E       F       G       H

SGTHAATAILINUIZBOISNIEDARSUEE
```

No. 2

```
  B       A       E       C

ERFOESMLIKEWBUDSGTHNATNEFNCAMY

  G       H       D       F

ZBOIHDIEDARSUTEINATANAAILIARI
```

A very important observation at the next step might easily be missed — although sometime in the subsequent steps you may have picked it up.

Let's assume you noticed right away that the IKE... sequences both "ended" with a D (if you consider as "ending" the last letter before another hit begins.) Similarly, the hits run ERFO..ML, with two letters hitting at the end, NEFN..MY, etc. — some ones, some twos finally — and occasional ambiguities which may or may not be accidental hits.

Even without taking a frequency count, the messages look like transpositions, right? And with eight areas of repetition, the hypothesis of a 8-wide box is reasonable. But why the hit-miss effect? (Or should one say hit-miss-hit?)

One gets such an effect when two columnar transposition messages, using different keys of the same width, start out alike, but have a somewhat different text shortly after. If it is merely a slight shift of the same text (such as an extra word in one message) the letter frequency would be substantially the same, and rather obviously so. Such is not the case here. Or is it?

We get a lot of the same letters appearing, but in different order. Let's list the ends of the columns — omitting the four-letter beginnings. (One uses the word "column" advisedly, since there seem to be 3 stretches of 8 letters and 5 stretches of 7 letters in our count of 59 letters each.)

Incidentally, the fact that they are the same length may be significant, since adding (or dropping) a word to account for the shift would make for different lengths.

```
A       B       C       D

IKEWRCDERFOADMLNEFNTHMYINATABA
```

```
  E       F       G       H
 ____    ____    ____    ____
SGTHAATAILINUIZBOISNIEDARSUEE

  B       A       E       C
 ____    ____    ____    ____
ERFOESMLIKEWBUDSGTHNATNEFNCAMY

  G       H       D       F
 ____    ____    ____    ____
ZBOIHDIEDARSUTEINATANAAILIARI
```

R	A	T	A	A	N	S	U	E	B	N	C	H	U	A	A
C	D	H	B	A	U	N	E	S	U	A	A	D	T	N	R
D	M	M	A	T	I	I	E	M	D	T	M	I	E	A	I
	L	Y				E		L			Y	E*			

*It is a simple matter to straighten out some of the seeming ambiguity of stretches, since we should end up with a transposition box having three long columns as the first three.

The ELY order of the long columns helps, but anyone can (immedately?) see the word IMMEDIATELY, and the problem is solved.

But one should also have observed that the letters on the RATA row of the first message fragment above occur for the most part on the SECOND row of the second fragment, while the CDHB letters on the next row hit best with the top row of the second set. And the total occurrences of the letters are matched exactly, box for box.

This happens to be a case of almost too many clues, but the point is made that such a "special solution" could have been reached without the "immediately" and without sweating over the "reading" of the message by matching good columns, and so forth.

All one would have to do is start with some column in Msg. #1, use the same "beginning" in Msg. #2 underneath it, then start chaining the letters in the "different" areas — by going from the fifth row of Box 1 to the sixth row of Box 2.

Can you see that since six out of the eight letters on row 5 of Box 1 go to Row 6 of Box 2, we should get a decimation of the order of at least six columns?

Try it anyway. Starting with a non-ambiguous letter such as R, we get R goes to U which goes to T which goes to A which (ambiguously) goes to S, N or A. But N goes to R and A goes to N, so we should have the best order as A N R U T (A) (S) — which has the interesting combination SATURNA in reverse.

When we "crib in" SATURNA on the top message, notice the fifth-to-sixth-row pairing:

```
S A T U R N A A
N D H E C U B A
```

This may sound complicated, but an alert cryppie could have used the SATURNA chaining approach, even without the IMMEDIATELY assistance. We emphasize this point because even in an unknown code or a fancy substitution (ADFGVX, for example) the "special solution" we have shown would put columns in a relatively correct order.

You can get the complete text now, by either route.

Perhaps several messages might be needed to iron out a relative solution, but you can see a hopeless task might be turned into a possible job because some code clerk — conscientiously using different transpositions keys — still gave the show away by saying ALMOST the same thing. (Saying the same thing with two keys wouldn't read the message — it would merely relate two keys.)

No. 22. You could have solved this by a number of different approaches, perhaps even without completely analyzing the system used.

Practiced analysts would have quickly said:

"Looks like a transpo, with X as word separator, and the lengths suggest a grille. Either some stupid clerk sent four parts of a message using the same size grille (and probably at the same setting) or four different (consecutive?) messages were similarly handled."

After a few minutes playing around on paper, the solver would have decided the latter situation prevailed, and had an answer in jig time.

The decision about sequent messages could have been amplified when the four messages were written out as a "depth" — since one would certainly want to compare the texts very closely.

```
NBOREXACFSOYERHXWEVC etc.
NVEREXXXREACILFONRGO
NSXRWXQXENRUIXYALONU
NDRRIXNMATIXEEUATNNI
```

Although there are many columns in the complete writeout (if the above comparison were continued) which have plenty of hits, the solid repeats around the first part are very interesting.

```
N R X
N R X
N R X
N R X
```

What more would you want? ("Nr." is often used for "Number" as opposed to "No." — removing any ambiguity about the negative.)

In any case, if the same position of a grille is reused, as seems to be the case here, there would be no trouble in multiple anagramming — even a couple of cases of QU are provided:

```
A Y  or  A C          O O
X C  or  X o          Y S
Q U              and  Q U
N X      N I          A N
```

You will note that the word separator is of doubtful help. There are just too many Xs at this point. But the NR beginning ought to yield something.

Ideally, we might find four consecutive numbers. What if we wrote the first letters of all consecutive numbers down, and looked to see if we could find a column in our 64-wide write out of the four messages which matched some portion of the series?

O T T F F S S E N T E T T F F S S E N T T T T T

Obviously if there were four messages in the fifties, or sixties, or whatever, we would get a solid column of F or S (or a satisfying combination) — so concentrate on the first part of the above.

Sure enough, soon after the X column, we get S E N T.

Looking for the columns which must go E I I E, and so on, is child's play. And since we are picking up approximately one out of four positions each time, it strengthens the idea of the grille, and all we have to do is "try it on the dog."

No. 23. Certain things are obvious, and others present phenomena that may be a trifle disguised. One can't miss the presence of only nine letters, which could imply some kind of a trinome system. But if you took my suggestion, and had a count of digraphs as well as one across the digraph cut, you could state that the digraphic count was overwhelmingly significant.

Whereas the cross-cut count has many doublets such as SS, DD, BB etc., the digraphic count only permits DD. And while letters seem to go across the cut with random association, the digraphic count shows that any given letter can go to a maximum of five other letters.

Moreover, there seem to be two classes —letters which have as partners A B D E or L and those letters which go to D O R S or W. And, of all things, this statement is true even if digraphs are reversed.

It seems apparent that we have a 5 x 5 square with a 5-letter keyword on the left and a different keyword on top — a slight variation on the usual Checkerboard, which in most cases has discrete representation of letters such as in BLACK and WHITE.

Your problem of finding something which made sense in the repeated letters of the crib is now answered, and reading the text should be a simple matter.

(Many of you probably got to the heart of the matter quickly by thinking about RBDBBDBR in the text.)

No. 24. This was too easy, but perhaps you will get in the habit of figuring simple odds for other things, even if you don't shoot craps.

The answer is obviously 1 out of 9 for getting your 9, but somewhat greater (6 out of 36 or 1 out of 6) of rolling the unlucky (or lucky if a first throw) 7. At the risk of insulting you, I'll list the ways: 4 5, 5 4, 3 6, 6 3 for the 9; and 7 comes from 3 4, 4 3, 5 2, 2 5, or 6 1, 1 6.

No. 25. This gave you an opportunity to see that cipher-text autokey is a lousy system. (Plain-text autokey is more secure, but subject to so many problems of error or garbling that it is not practical.)

You probably noticed several phenomena. There is not enough text to get terribly excited about frequencies, but the repeat is certainly significant if you took an arithmetical difference (in a consistent direction, of course) between letters, assuming a normal alphabet.

Such delta counts are seldom taken or such phenomena found, but you never can tell.

If you wrote out the delta stream that way, it would read:

```
8 15 7 4 17 19 4 23 19 0 20 19 14 10 4 24 8 18 13 14 19
16 19 8 19 4 18 14 5 14 20 11 4 3 20 15 1 24 4 17 17 14
17 18 0 18 15 11 0 8 13 19 4 23 19 0 20 19 14 10 4 24
```

The repeat of 0 20 19 14 10 4 24 is undoubtedly significant. What does it mean? First — and most important — it proves that normal alphabets were used in the encipherment. (Obviously, we would have gotten garbage by this procedure if we had used wrong rules for differences.)

The repeat starts at the 10th position and the 56th, which gives a possible cycle as a factor of 46, which is hardly exciting. However, the numbers themselves are rough — extremely so:

```
 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20
 1     8 1   1 4    2  2     2  6  3  1  4  4  9  4

21 22 23 24 25 0
      2  3    3
```

Experienced cryppies would immediately recognize the frequency of the normal alphabet, 4 being E, 19 being T, and so on. This smacks of an ordinary Vigenere encipherment, wherein A key makes A plain equal to A cipher. And since the numbers came from the delta stream, it necessitates a cipher-text autokey.

A plain-text autokey has properties which are rather different. If there is a repeat in the plain, it will show up as a repeat in the cipher, shorter by the amount of offset in the beginning of the plain key. If you had agreed to use "A" (or X or any other letter) as an introductory key to encipher the first plain, and then use the first plain to encipher the second, the offset of one would make the repeat one letter "shorter."

An arranged 5-wide starting key would cut down the length of any meaningful repeats by five letters, which would not help suppress significant repeats.

But since you don't have the tool of the key placed right there for you to see, as in cipher-text autokey, the going is much rougher in plain-text auto-key.

I just wrote out a similar plain message and enciphered it by plain-text autokey with one introductory letter. The cipher is flat — starting out:

PALIV GXBQT UNHYO CFALH CRDNT KVEIO etc.

But because the offset was only one, you could easily read this by assuming the starting K of A, B, C in turn — if you were lucky enough to have Vigenere encipherment, or knew the alphabets in some other arrangement. With an introductory key of, say, 5 wide, you could make a trial test which would give you a string of plain values in the 1st, 6th, 11th position, etc. Then you would have to test a string of the 2nd, 7th, 12th, and so on.

Theoretically, if the text was long enough, you could decide which of 26 possible strings were statistically "plain" enough at each of the five offset positions, but I wouldn't tackle such a job unless the reward were more important than the satisfaction gained by solving these puzzles.

You would have to guess the method of encipherment, the length of introductory key and do a lot of pencil-pushing on a hunch.

No. 26. So we gave you something we didn't intend to. The system is a dictionary code, with variants on the values which stand for page numbers and word separators.

It's not all that easy to read such a single message even after breaking the method of enciphering the code values. We will mention how to interpolate values, but your best chance (if it were really an enemy spy message or something) might be to spend a few hours —or days — tracking down a possible dictionary — for there is a strong hint that the letter A is on page ll, line 1. (No values precede this, and most dictionaries have a lot of junk on the first few pages before definititions start — while a real code would have little reason to start out on page 11.)

Of course, "breaking" a one-part code is fairly straightforward if you have plenty of traffic — since you know the plain values probably run in order, something like A ABLE ACT ACTION... to ZONAL ZONE. If there are special areas for punctuation, numbers, and other handy charts, such things will come out in the wash.

A two-part code comes from a different fish kettle, and we won't even think about it here. Now what did you list as "monkey business?"

First of all, even before taking a count, you no doubt saw the striking absence of repeated letters as digraphs — I think the ZZ is the only occurrence.

And you may have noticed the peculiar rhythm of certain letters, such as the last letters of the alphabet. Starting with the fourth letter U, there is an obvious pattern of the six finals, as we see a V after a few positions, and then a W, then an X and Y and Z—and bingo. We're back to V W X Y Z — but then the parade gets ragged.

This might lead you to ask yourself, "Do other letters behave like that?" It starts out interestingly as A B C — but peters out. However, A and B have a sort of rhythm.

How about the frequency count you took? Amazing. (My figures may have missed a letter or so, but:

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
27	19	22	18	17	18	13	11	11	13	10	5	11	5	5	4	3	5	3	3

and then a leap to high frequencies of 13 for each of the letters U to Z.

Moreover, with these last letters of the alphabet, the counts are no more exciting than the specifics of what follows — in practically all cases, the last six letters of the alphabet are followed by letters included in the first twenty alphabetical positions.

What can cause such a weird distribution? Both the rhythm and the matching (more or less) frequencies of pairs of letters suggest the use of variants. As we have said, such attempt at suppressing repeats by occasionally alternating variants on a predictable basis might backfire, as it does in this case.

And the frequency of the U to Z distribution suggests that these letters are being used as word separators — a theory quite consistent with having 20 letters, in variant pairs, to represent 10 numbers.

When the assumption is made that A and B represent 1, C and D stand for 2, and so on, the message obviously breaks down into a maximum of a three-digit page and a maximum of a 2-digit line number. This has all the earmarks of a dictionary code, in which page numbers and line positions are limited.

An ordinary one-part code as used by serious communicants (supposedly enciphered by some means, such as an additive, to offer some degree of security) would normally be in four- or five-digit groups (or equivalent letter groups) running from perhaps 00001 or 0001 to however high a number was necessary to accommodate the desired vocabulary. There would be no reason to suppose that one digit had a much greater frequency than any other.

In the case we are examining, a proper digital write-out would be:

```
465-1  133-32  94-12  251-23  379-6  11-1  321-4
333-43 94-12  (Two separators standing for period?)
252-6 251-23 312-16 and so on.
```

Note that the pages have as highest value 522, with two words close together. (YOU and YOUR perhaps?) The highest line number (or position on page, since most dictionaries have two columns on each page) seems to be 43, which is reasonable.

Naturally, this causes the frequency of the higher numbers to be much smaller than the lower numbers (which can be used in any position). This shows up very clearly when counts are taken by position.

Why did we start by assuming A/B was 1 rather than 0? Because there are single digit line numbers — 1 and 6, for instance, instead of 01 and 06. (Why waste a space, especially because the variation in length of page or line number eliminates any possible cycle of five values per group.)

How do we go about solving such a critter? A crib would help — suppose I gave you the start of "The Dictionary Code is really a" — you would have a few anchor points to work from.

Of course, repetitions (such as the beginning THE and the A) might lead you to try common words in the general area of expected occurrence in a dictionary.

In "real life" one might expect spellers, since that proper word you have to use, such as HITLER, might not be in the dictionary your are using.

Spellers would either start with a special group (and there could easily be a page of such, with END SPELL as well) or the ordinary group for SPELL. Then one would get the groups for H, I, T and so forth.

You can see that such alphabetical spellers would soon be identified if one had sufficient traffic, since they would cluster, and have monoalphabetic frequency. Also names such as MORROW give nice patterns.

If numbers are used in the text, such as "725 81st AVE" they of course would have to be words — NINE TWO FIVE — and they, too, would cluster. (A special page with groups that yielded such as "01-17 01-12 01-15 01-18" etc., would make our job even easier.)

With enough traffic, breaking a one-part code is no big deal. On a few messages, suspecting a dictionary, one might use interpolation to help out.

This is done by using ANY dictionary of approximately the right size, and by assuming such things as A, THE, and YOUR in the message we gave you, get a few anchor points on the comparative listing. To prove my point, I am deliberately using a dictionary that is much larger, with a wide disparity between the highest and lowest matches: This is all theoretical, so we needn't worry about how we got the idea that 282-34 is MANY. Thus:

Message group	Meaning	Dictionary group in trial
11-1	A	1-1
94-12	CODE	138-22
251-23	IS	389-13
282-34	MANY	437-49
465-1	THE	710-43
522-15	YOU	797-49
522-18	YOUR	798-6

You will see that the YOU YOUR stretches imply more words in the larger dictionary, naturally.

More talented mathematicians than I could show you exactly what formulae to apply to bring these two stretches into line (11 to 522 and 1 to 798, since we need only work with pages.) But you can see that two comparative lines would read:

```
11.....94.............251...282.........465.....522
 1.....138............389...437.........710.....797
-10      +42             +148  +155         +255     +265
```

Even crude interpolation might help. When I fiddled around with my "big" dictionary numbers, I came up with the seventh and eighth words as being around the start of the Ps.

Remember we can be quite a bit off, until we get more anchor points, so the two consecutive groups could start with an O or a P group followed by another O or P group — or an unlikely Q group. Even without my prodding, you might have logically decided on ONE PART.

You can see that sooner or later, and with more traffic and better anchor points — or using a more compatible dictionary than my deliberate misfit — you would hit on THE DICTIONARY CODE IS REALLY A ONE PART CODE, IT IS NO DIFFERENT, and so on to the sweet end.

But remember the reading of the message was not the main idea — it was the finding and explaining of the phenomena at which I was aiming.

Having given the diagnostic approach a boost (I hope) over the tell-me-and-I'll-do-it type of puzzle conquering, I will leave you with the words I often see at the bottom of letters and articles:

Good Solving.

Index

A

ACA
    See American Cryptogram Association
ADFGVX, 2, 147, 224, 239
Aegean Park Press, 147
Affinity of letters, 22
Alphabets (Reconstruction), 35, 163, 169
The American Black Chamber, 8-9
American Cryptogram Association, 1-3, 6-7, 11-13, 15, 63, 92, 94, 100, 115, 140, 189, 191, 233
The American Magic, 8
Amsco, 209, 211
Anagrams, 220, 228
Aristocrats
    See Monoalphabetic Substitution with Word Separator
Arlington Hall, 8
Autokey, 120-121, 181, 241-242

B

Francis Bacon, 126
Baconian cipher, 113, 126
James Bamford, 127
Beaufort, 113, 119-120
Bifid, 69-70, 76, 82, 87, 103, 105, 163, 181, 223, 234-235
    C.M. Bifid, 83
    5 x 5 Bifid, 80
    6 x 6 Bifid, 78, 80
    Twin Bifid, 75
Blackjack, 151, 161
Bletchley Park, 5, 7, 42
Bodyguard of Lies, 8
Book Cipher, 122-123, 222
Boustrophedon, 67, 138, 163-164
British crossword
    See cryptic crossword puzzles

C

Caesar slide, 11-12, 24, 119, 125
Lambros Demetrios Callimahos, 86
Peter Calvocoressi, 42
Card drawing, 156
Anthony Cave Brown, 8

Checkerboard, 40, 241
Ronald Clark, 7
C.M. Bifid
    See Congugated Matrix Bifid
The Codebreakers, 7, 122, 127, 158
Computers, 1, 33, 189, 216
Conjugated Matrix Bifid, 83
Constant Universe, 150
Crib, placing of, 128
The Cryptogram, 2-3, 6, 191
Cryptic crossword puzzles, 2-3, 213-214
Cryptologia, 3
Cycle, 20, 43

D

Decimation, 20, 29-30, 38, 59, 61, 170, 238
Delta, 161, 241-242
Paul Derthick, 27, 221, 231
Diagnosis, 42, 175, 189, 213, 223-224
Diagnostics, 161, 177, 189
Dice, 149-150, 225
Dictionary code, 186, 242-243
Digraphic substitution, 1, 230
Digraphid, 87
Direct Symmetry
    See Symmetry
Distribution, 179

E

Enigma, 2, 5-8, 41-42, 111, 121, 174, 179, 232
Explaning the Phenomona, 179
Exploitation, 8, 174, 176-177, 193

F

Col. Fabyan, 7, 126
5-5-3 Treaty, 8
Foreign Language Ciphers, 26, 31, 223, 232
Foursquare, 141
Fractionated Morse, 100-101, 205
Fractionation, 68-69, 86-87, 100, 103, 224
French monoalphabet, 26
Frequency Count, 1, 17, 21, 25, 36, 111, 130, 178-179, 182, 184, 188-189, 222, 228, 231, 237, 243
Elizebeth Friedman, 126
William Friedman, 3, 5-9, 51, 86, 122, 126-127, 147, 158, 169, 175, 230

G

GC&CS
See Bletchley Park
German ciphers
See Enigma
Grandpre, 113, 202
Grille, 41, 126, 138-145, 232, 239-240
Gromark, 116, 126
Gronsfeld, 114-116

H

Hatting, 135, 137
Headline Puzzles, 27, 231
Hindus, 230
Homophonic, 201, 236
The Hut Six Story, 6, 8

I

I.C.
See Index of Coincidence
Identity (in Phenomena), 177-178, 184
Index of Coincidence, 157, 190
Indirect Symmetry
See Symmetry
Intrepid, 7

J

Japanese Naval Conference
See 5-5-3 Treaty

K

K-1, 13, 21, 35, 169, 233
K-2, 15, 169, 233
K-3, 28, 83, 170
K-4, 83, 170
David Kahn, 7, 126
Kappa value, 159
Keywords, 15-16, 20, 23-25, 27, 31-32, 34-35, 37-40, 43, 47, 50, 54-55, 60, 62-63, 67, 79, 85-87, 93, 96, 99, 101-102, 104, 107-108, 110, 119-120, 123, 125, 129, 131-132, 163-170, 185-186, 201, 205, 208, 210, 220, 227-228, 233, 241
Kryha, 125-126
Solomon Kullback, 9, 127, 217

L

John Larkin, 127
Latent phenomena, 178, 180-181
Latin squares, 217-218
Ronald Lewin, 8

M

The Man Called Intrepid, 7
Martin and Mitchell, 127
Mixed alphabets, 12, 170
Monoalphabetic Substitution with Word Separator, 2,15, 33, 36, 214
Monoalphabetic Substitution Without Word Separator, 2, 15, 33, 36, 214
Morbit, 205, 208, 218
Myszkowski, 210

N

The Nation, 1-3, 160, 214, 220
NSAers, 127
Numerical key recovery, 177

O

Obvious phenomena, 180
Odds, 72, 150-155, 158, 160, 179, 241

P

Patristocrats
    See Monoalphabetic Substitution Without Word Separators
Pattern
    See word pattern
Pearl Harbor investigation, 5
Joseph Peterson, 127
Phenomena, 177, 179
Pig-pen Cipher, 39
Playfair, 1-2, 6, 14, 33, 63-65, 67-68, 113, 123, 125, 163, 172, 174, 179-180, 183, 189, 228
Pollux, 205, 218
Polyalphabetic substitution
    See Quagmire
Porta, 115, 127
Portuguese monoalphabet, 26
Probability, 106, 111, 149-150, 153, 155, 157, 180, 219
Problem Solving, 1-2, 111, 213
Problems to Solve, 31, 120, 191, 219, 227
Purple, 5-7

Q

Quagmire, 43, 128, 180, 191, 193, 202, 223-224, 234
Q-I, 43-44
Q-II, 43, 47
Q-III, 43, 51
Q-IV, 55

R

Railfence, 197
Random, 115, 153, 159, 161, 218
Relationship, 179
Replacement, 149-151, 155, 157-158, 160
Riverbank Publications, 147
Roughness, 48, 158, 172, 174, 179, 181, 226
Roulette, 150-152, 155
Frank Rowlett, 9

S

Scrabble, 149, 160, 216
Seriated Playfair, 123, 128
Shakespeare, 126
The Shakespearean Ciphers Examined, 126
Sherlock Holmes, 41
Abraham Sinkov, 6, 9, 45, 158
6 x 6 Bifid
    See Bifid
Spycatcher, 8
Squares and square recovery, 86, 163-164
Statistics, 149
William Stevenson
    See *The Man Called Intrepid*, 7
Symmetry, 29, 53, 55

T

Margaret Thatcher, 8
Tip
    See Crib
Transposition, 7, 41-42, 132, 135, 138, 181, 197-198, 212
    Completely filled columnar, 210-211
    Incompletely filled columnar, 209
    Grille, 41, 126, 138-145, 232, 239-240
    Route Transposition, 137
    Swagman, 198, 217-218, 233
Trifid, 103, 105, 113
Tri-Square, 108

Twin Bifid, 75, 81-82
Two-Square, 97

V

Variant, 119-120
Col. Verkuyl, 127
Vigenere, 119-121, 126-127, 242

W

Eugene Waltz, 125
Gordon Welchman, 8
Winds messages, 5
Word pattern, 14-15

X

Xenocrypt
See Foreign Language ciphers

Y

Herbert O. Yardley, 3, 6, 8-9, 127

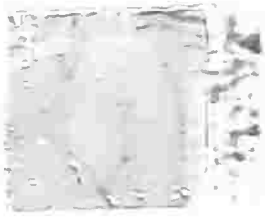
Z

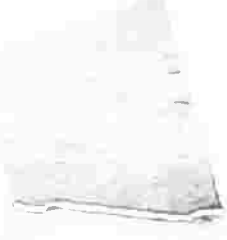
The Zimmermann Telegram, 41

BIBLIOGRAPHY

American Cryptogram Association, *The ACA and You*, ACA, 1988
Barker, Wayne G., *Cryptanalysis of the Hagelin Cryptograph*, Aegean Park Press, 1980
Beesly, Patrick, *Very Special Intelligence*, Ballantine Books, 1981
Callimahos, Lambros D., *Military Cryptanalytics, Part I*, Aegean Park Press, 1985
Calvocoressi, Peter, *Top Secret Ultra*, Ballantine Books, 1981
Carlisle, Sheila, *Pattern Words: Three Letters to Eight Letters*, Aegean Park Press, 1986
Cave Brown, Anthony, *Bodyguard of Lies*, Bantam Books, 1977
Clark, Ronald, *The Man Who Broke Purple*, Little, Brown & Co., 1977
Farago, Ladislas, *The Broken Seal*, Random House, 1967
Foster, Caxton, *Cryptanalysis for Microcomputers*, Hayden, 1982
Friedman, William F., *Elements of Cryptanalysis*, Aegean Park Press, 1979
Gaines, Helen Fouche, *Cryptanalysis*, Dover, 1956,
Glover, D. Beaird, *Secret Ciphers of the 1876 Presidential Election*, Aegean Park Press, 1991
Hilton, Stanley E., *Hitler's Secret War in South America*, Ballantine Books, 1981
Hodges, Andrew, *Alan Turing, The Enigma of Intelligence*, Unwin Hymn, Ltd., 1985
Kahn, David, *The Codebreakers*, Macmillan Co., 1967
Kullback, Solomon, *Statistical Methods in Cryptanalysis*, Aegean Park Press, 1977
Lewin, Ronald, *American Magic*, Penguin Books, 1982
Millikin, Donald D., *Elementary Cryptography and Cryptanalysis*, Aegean Park Press, 1991
Partridge, Eric, *Origins*, Macmillan Co.
Sinkov, Abraham, *Elementary Cryptanalysis*, Yale University, 1966
Stevenson, William, *A Man Called Intrepid*, Book Club Associates, 1976
Tuchman, Barbara W., *The Zimmermann Telegram*, Macmillan, 1958
Welchman, Gordon, *The Hut Six Story*, Allen Lane, 1982
West, Nigel, *G. C. H. Q.*, Coronet Books, 1987
Winterbotham, F. W., *The Ultra Secret*, Weidenfeld and Nicolson, 1974
Yardley, Herbert O., *The American Black Chamber*, Aegean Park Press, 1987

Note: This bibliography is a small representation of the many books concerning cryptography and cryptanalysis. For the serious historian or student looking for additional titles, a good source is *Books in Print* which can be found in most libraries and bookstores. For further titles, consult the catalog of Aegean Park Press.





For thirty years Frank W. Lewis actively engaged as a senior cryptanalyst with the United States Government. He began his career in 1939 with the Signal Intelligence Service when he was hired by Mr. William Friedman (*The Man Who Broke Purple*) as one of the especially qualified person to whom the burden would fall (to read an enemy's secret messages) in what Friedman knew would be inevitable hostilities. That Friedman had keen insight and judgement to select the right people is evidenced not only by his selection of the already publicized triumvirate — Frank B. Rowlett, Abraham Sinkow, and Solomon Kullback — but also by his selection of Frank W. Lewis. While much of the work of Frank W. Lewis must remain forever behind a veil of secrecy, this much can be said: Lewis may be the only person who has received the War Department's Medal for Exceptional Civilian Service and a similar medal from the National Security Agency.

As far as the other not-quite-so-exotic fields are concerned, since 1947 *The Nation* magazine has featured the very challenging "Cryptic" crossword puzzle of Frank W. Lewis each week. Although not many puzzlers on this side of the ocean were acquainted with this British import, a few aficionados have always rallied to his cause, including some authoritative voices. *Harper's* in 1950 published an article in which they names their candidates for the four outstanding puzzlers: Ronald Carton of the London *Times*, Ruth Brooks of the Boston *Herald* for her "straight" crosswords, Elizabeth Kingsley of Doublecrostic fame, and Frank W. Lewis.

Times Books has initiated a series of eight volumes of Frank's *The Nation* puzzles, with one book to appear every six months. This greatly expanded exposure perhaps marks the end of the period when *The Nation* puzzles were unknown except to a small but vociferous group of loyal fans, and implies a growing sophistication in American puzzle tastes.

Except for a period of consultative service, Lewis retired from NSA in 1969 and moved with his wife, Sylvia, to the Caribbean paradise of Montserrat, where he is still adding to the stack of a few thousand *The Nation* tear-sheets. The couple celebrated their 55th wedding anniversary in 1991 in a get-together with their five children, who are scattered around the globe.

Frank's original field was English, but his master of music degree (on top of a bachelor's degree and a couple of years each of business college and teachers college training) somehow met the requirements for the higher professional ranks within the Civil Service. This melange of training and interest stood him in good stead both as a cryptanalyst and as a retired life-liver, where the computer is competing with bridge and cryptograms for his seemingly boundless enthusiasm.

ISBN 0-89412-178-2

90000

9 780894 121784

מכללת אורנים ע"ר
מספרים
003352200425
SOLVING CIPHER PROBLEMS